RESTRICTION OF LAPLACE-BELTRAMI EIGENFUNCTIONS TO CANTOR-TYPE SETS ON MANIFOLDS

SURESH ESWARATHASAN AND MALABIKA PRAMANIK

Abstract. Given a compact Riemannian manifold (M, g) without boundary, we estimate the Lebesgue norm of Laplace-Beltrami eigenfunctions when restricted to certain fractal subsets Γ of M. The sets Γ that we consider are random and of Cantor-type. For large Lebesgue exponents p, our estimates give a natural generalization of L^p bounds previously obtained in [16, 17, 33, 8]. The estimates are shown to be sharp in this range.

The novelty of our approach is the combination of techniques from geometric measure theory with well-known tools from harmonic and microlocal analysis. Random Cantor sets have appeared in a variety of contexts before, specifically in fractal geometry, multiscale analysis, additive combinatorics and fractal percolation [19, 21, 22, 29, 30]. They play a significant role in the study of optimal decay rates of Fourier transforms of measures, and in the identification of sets with arithmetic and geometric structures. Our methods, though inspired by earlier work, are not Fourier-analytic in nature.

Contents

Date: January 20, 2019.

2010 *Mathematics Subject Classification.* 28A33, 28A78, 28A80, 35J05, 35P20, 42B37, 47A60, 47A75, 47J10, 58J05 (primary), and 26A33 (secondary).

1. Introduction

The study of eigenfunctions of Laplacians lies at the interface of several areas of mathematics, including analysis, geometry, mathematical physics and number theory. These special functions arise in physics and in partial differential equations as modes of periodic vibration of drums and membranes. In quantum mechanics, they represent the stationary energy states of a free quantum particle on a Riemannian manifold.

Let (M, g) denote a compact, connected, n-dimensional Riemannian manifold without boundary. The ubiquitous (positive) Laplace-Beltrami operator on M, denoted $-\Delta_g$, is the primary focus of this article. It is well-known [34, Chapter 3] that the spectrum of this operator is

non-negative and discrete. Let us denote its eigenvalues by $\{\lambda_j^2 : j \geq 0\}$, and the corresponding eigenspaces by $\mathbb{E}_j$. Without loss of generality, the positive square roots of the distinct eigenvalues can be arranged in increasing order, with

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots \lambda_j < \cdots \to \infty.$$

It is a standard fact [34, Chapter 3] that each $\mathbb{E}_j$ is finite-dimensional. Further, the space $L^2(M, dV_g)$ (of functions on M that are square-integrable with respect to the canonical volume measure dV_g) admits an orthogonal decomposition in terms of $\mathbb{E}_j$:

$$L^2(M, dV_g) = \bigoplus_{j=0}^{\infty} \mathbb{E}_j.$$

One of the fundamental questions surrounding Laplace-Beltrami eigenfunctions targets their concentration phenomena, via high-energy asymptotics or high-frequency behaviour. There are many avenues for this study. Semiclassical Wigner measures provide one way to measure concentration, as exemplified in the seminal work of Shnirelman [31], Zelditch [38], Colin de Vèrdiere [9], Gérard and Leichtnam [12], Zelditch and Zworski [39], Helffer, Martinez and Robert [14], Rudnick and Sarnak [27, 28], Lindenstrauss [23], and Anantharaman [1]. Another direction involves growth of the L^p norms of these eigenfunctions. The contribution of this article lies in the latter category. Specifically, it describes the $L^2(M) \to L^p(\Gamma)$ mapping property of a certain spectral projector (according to the spectral decomposition above), where Γ is a fractal-type subset of M. In particular, Γ does not enjoy any smooth structure, a point of departure from prior work where this feature was heavily exploited. We begin by reviewing the current research landscape that will help place the main result Theorem 1.3 in context.

1.1. **Literature review.** The Weyl law, itself a major topic in spectral theory, provides an L^∞ bound on eigenfunctions on M [16]. The first results that establish L^p eigenfunction bounds for $p < \infty$ are due to Sogge [33].

Theorem 1.1. [33] *Given any manifold M as above and $p \in [2, \infty]$, there exists a constant $C = C(M, p) > 0$ such that the following inequality holds for all $\lambda \geq 1$:*

$$\|\varphi_\lambda\|_{L^p(M)} \leq C(1+\lambda)^{\delta(n,p)} \|\varphi_\lambda\|_{L^2(M)}, \text{ with}$$

$$\delta(n,p) = \left\{ \begin{array}{ll} \dfrac{n-1}{4} - \dfrac{n-1}{2p}, & \text{if } 2 \leq p \leq \dfrac{2(n+1)}{n-1}, \\ \dfrac{n-1}{2} - \dfrac{n}{p}, & \text{if } \dfrac{2(n+1)}{n-1} \leq p \leq \infty. \end{array} \right\} \tag{1.1}$$

Here φ_λ is any eigenfunction of $-\Delta_g$ corresponding to the eigenvalue λ^2. The bound is sharp for the n-dimensional unit sphere $M = \mathbb{S}^n$, equipped with the surface measure.

Historically, an important motivation and source of inspiration for this line of investigation has been the Fourier restriction problem, which explores the behaviour of the Fourier transform when restricted to curved surfaces in Euclidean spaces. In fact the Stein-Tomas L^2 restriction theorem [36], originating in Euclidean harmonic analysis, was a key ingredient in an early proof of Theorem 1.1 for the sphere. Indeed, Theorem 1.1 may be viewed as a form of discrete restriction on M where the frequencies are given by the spectrum of the

manifold, see for example [32]. Conversely, it is possible to recover the L^2 restriction theorem for the sphere from a spectral projection theorem such as Theorem 1.1 applied to the n-dimensional flat torus. The lecture notes of Yung [37, Section 2] contain a discussion of these implications.

Theorem 1.1 permits a number of independent proofs. For an argument that involves well-known oscillatory integral estimates of Hörmander applied to the smooth spectral projector (denoted $\rho(\lambda - \sqrt{-\Delta_g})$), we refer the reader to the treatise [34]. The semi-classical approach of Koch, Tataru and Zworski [20] has also yielded many powerful applications.

Finer information on eigenfunction growth may be obtained through L^p bounds on φ_λ when restricted to smooth submanifolds of M. One expects φ_λ to assume large values on small sets. Thus its L^p-norm on a Lebesgue-null set such as a submanifold, if meaningful, is typically expected to be larger in comparison with the L^p norm taken over the entire manifold M, as given by Theorem 1.1. The first step in this direction is due to Reznikov [26], who studied eigenfunction restriction phenomena on hyperbolic surfaces via representation theoretic tools. The most general results to date on restricted norms of Laplace eigenfunctions are by Burq, Gérard and Tzvetkov [8], and independently by Hu [18]. The work of Tacy [35] has extended these results to the setting of a semi-classical pseudo-differential operator (not merely the Laplacian) on a Riemannian manifold, while removing logarithmic losses at a critical threshold. Another particular endpoint result is due to Chen and Sogge [10]. We have summarized below the currently known best eigenfunction restriction estimates for a general manifold, combined from this body of work and for easy referencing later.

Theorem 1.2. [8, 18, 35] *Let $\Sigma \subset M$ be a smooth d-dimensional submanifold of M, equipped with the canonical measure $d\sigma$ that is naturally obtained from the metric g. Then for each $p \in [2, \infty]$, there exists a constant $C = C(M, \Sigma, p) > 0$ such that for any $\lambda \geq 1$ and any Laplace eigenfunction φ_λ associated with the eigenvalue λ^2, the following estimate holds:*

$$\|\varphi_\lambda\|_{L^p(\Sigma, d\sigma)} \leq C\,(1+\lambda)^{\delta(n,d,p)} \|\varphi_\lambda\|_{L^2(M, dV_g)}. \tag{1.2}$$

The exponent $\delta(n, d, p)$ admits a multi-part description. Specifically,

$$\delta(n, n-1, p) = \left\{ \begin{array}{ll} \dfrac{n-1}{4} - \dfrac{n-2}{2p}, & \text{for } 2 \leq p \leq \dfrac{2n}{n-1}, \\ \dfrac{n-1}{2} - \dfrac{n-1}{p}, & \text{for } \dfrac{2n}{n-1} \leq p \leq \infty. \end{array} \right\} \tag{1.3}$$

For $d \neq n-1$,

$$\delta(n, d, p) = \frac{n-1}{2} - \frac{d}{p}, \qquad \text{for } 2 \leq p \leq \infty \text{ and } (d,p) \neq (n-2, 2). \tag{1.4}$$

For $(d, p) = (n-2, 2)$, the exponent $\delta(n, d, p)$ is still given by (1.4); however, there is an additional logarithmic factor $\log^{1/2}(\lambda)$ appearing in the right hand side of inequality (1.2).

The proofs in [8] and [10] use a delicate analysis of oscillatory representations of the smoothed spectral projector $\rho(\lambda - \sqrt{-\Delta_g})$ restricted to submanifolds Σ, combined with refined estimates influenced by the considered geometry. Alternatively, [18] uses general mapping properties for Fourier integral operators with prescribed degenerate canonical relations to obtain bounds for the oscillatory integral operators in question. There are several recurrent features

in these proofs; namely, stationary phase methods, arguments involving integration by parts, operator-theoretic convolution inequalities. This methodology heavily relies on the fact that the underlying measures are induced by Lebesgue, which in turn is a consequence of M and Σ being smooth manifolds. The present article explores the accessibility of this machinery in the absence of smoothness, and aims to find working substitutes when such methods are unavailable. This leads to a discussion of our main results.

1.2. **Main results.** An interesting feature of the exponents $\delta(n,p)$ and $\delta(n,d,p)$ occurring in Theorems 1.1 and 1.2 respectively is that for large p, they are both of the form $(n-1)/2-\alpha/p$, where

$$\begin{aligned} \alpha &= \text{dimension of the space on which the } L^p \text{ norm of } \varphi_\lambda \text{ is measured} \\ &= \left\{ \begin{array}{ll} \dim(M) = n & \text{in Theorem 1.1,} \\ \dim(\Sigma) = d & \text{in Theorem 1.2.} \end{array} \right\} \end{aligned} \tag{1.5}$$

In view of this commonality in (1.1), (1.3) and (1.4), we pose the following question: is there a class of "sparser" sets $\Gamma \subseteq \Sigma$, or equivalently a class of measures μ that are singular relative to the canonical measure on Σ, with respect to which we can estimate the growth of our eigenfunctions φ_λ? The optimal scenario would be to obtain bounds that reflect the dimensionality of the set Γ in the same way that Theorems 1.1 and 1.2 do. We answer this by presenting the main result of our article.

Theorem 1.3. *Fix positive integers $n \geq 2$ and $1 \leq d \leq n$. Let Σ be a smooth, d-dimensional submanifold of M. For each $\varepsilon \in [0,1)$, we define the critical exponent*

$$p_0 = p_0(n,d,\varepsilon) := \frac{4d(1-\varepsilon)}{n-1}. \tag{1.6}$$

Then for each choice of n, d, Σ and ε, there is a probability space $(\Omega, \mathcal{B}, \mathbb{P}^*)$ depending on these parameters that obeys the properties listed below.

(a) For $\mathbb{P}^*$-almost every $\omega \in \Omega$ there exists a Cantor-type subset $\Gamma_\omega \subset \Sigma$, equipped with a natural probability measure ν_ω, such that the set Γ_ω has Hausdorff dimension $d(1-\varepsilon)$. For $\varepsilon = 0$, ν_ω is singular with respect to the natural surface measure on Σ induced by the Riemannian metric g.

(b) For $\mathbb{P}^*$-almost every set Γ_ω obtained in (a) *there exists a finite constant $C = C(\omega, n, d, p, \varepsilon) > 0$ such that for all $\lambda \geq 1$, we have the eigenfunction estimate*

$$\|\varphi_\lambda\|_{L^p(\Gamma_\omega,\nu_\omega)} \leq C\lambda^{\delta_p}\, \Phi(\lambda)\, \|\varphi_\lambda\|_{L^2(M,dV_g)}. \tag{1.7}$$

Here φ_λ denotes any L^2-eigenfunction associated with the eigenvalue λ^2 for the Laplace-Beltrami operator $-\Delta_g$ on M. For $p_0 > 2$, the exponent δ_p is given by

$$\delta_p = \delta_p(n,d,\varepsilon) := \left\{ \begin{array}{ll} \dfrac{n-1}{4}, & \text{if } 2 \leq p \leq p_0, \\ \dfrac{n-1}{2} - \dfrac{d(1-\varepsilon)}{p}, & \text{if } p \geq p_0. \end{array} \right\} \tag{1.8}$$

For $p_0 \leq 2$, the exponent $\delta_p = (n-1)/2 - d(1-\varepsilon)/p$ for $2 \leq p \leq \infty$. The quantity $\Phi(\lambda)$ appearing in (1.7) *is an increasing function that grows slower than any positive power of*

λ; specifically, Φ is of the form

$$\Phi(\lambda) = \exp(C'\sqrt{\log(\lambda)}), \tag{1.9}$$

where $C' = C'(n, d, p, \varepsilon) > 0$ is an explicit constant.

(c) The exponent δ_p in the above estimate is sharp in general for $p \geq \max(p_0, 2)$, in the following sense. Suppose that Σ is any d-dimensional submanifold of the n-dimensional unit sphere $M = \mathbb{S}^n$, $d \leq n$. Fix $\varepsilon \in (0, 1]$.

There exists a sequence of L^2-normalized spherical harmonics $\{\varphi_{\lambda_j} : j \geq 1\}$ with $\lambda_j \nearrow \infty$ such that for $\mathbb{P}^*$-almost every set Γ_ω obtained above and for every $p \geq p_0$, one can find a constant $C = C(\omega, p) > 0$ verifying the lower bound

$$\|\varphi_{\lambda_j}\|_{L^p(\Gamma_\omega, \nu_\omega)} \geq C\lambda_j^{\delta_p}\Phi(\lambda_j)^{-1} \tag{1.10}$$

for all λ_j sufficiently large.

1.3. **Remarks.** Let us pause for a moment to contextualize some of the important features of our result, and expand on directions of further improvement.

1. For $p \geq p_0$, the exponent δ_p in Theorem 1.3 (b) is of the same form alluded to in (1.5), namely $\delta_p = (n-1)/2 - \alpha/p$ with $\alpha = d(1-\varepsilon)$. Thus our result may be viewed as a natural interpolation between the global estimates in [33] and the smooth restriction estimates in [8], bridging the estimates across a family of sets with continuously varying Hausdorff dimensions.

2. To the best of our knowledge, Theorem 1.3 is the first result of its kind in several distinct categories. First, it offers, for *every manifold M and every smooth submanifold Σ therein*, eigenfunction bounds over non-smooth subsets of positive but non-integral Hausdorff dimension. Second, even for integers m, our result produces new sets of dimension m, for example with $(n, d, \varepsilon) = (2, 2, 1/2)$, that are not necessarily contained in any m-dimensional submanifold, and yet capture the same eigenfunction growth bounds as smooth submanifolds of the same dimension, up to sub-polynomial losses. Third, when $\varepsilon = 0$, our result provides examples of singular measures supported on submanifolds with respect to which the eigenfunctions obey the same L^p growth bounds as with the induced Lebesgue measure on the same submanifold. This is reminiscent of an earlier article by Łaba and Pramanik [22], where the authors construct a random Cantor-type measure with respect to which the maximal averaging operator has the same L^p mapping properties as the Hardy-Littlewood maximal function (where the underlying measure is Lebesgue).

3. As in [8, 18], the proof of Theorem 1.3 yields estimates not merely for the eigenfunctions φ_λ, but also for the smoothed spectral projector on M. The sharpness statement for $p \geq p_0$ continues to hold for such operators.

4. On the other hand, Theorem 1.3 leaves room for improvement on several fronts. As we will explain in the next section, our methods may be viewed as a fractal adaptation of [8, Theorem 1], which itself ignores oscillations inherent in an underlying operator. While this yields sharp results (excluding the $\Phi(\lambda)$ factor) in the full range $2 \leq p \leq \infty$ if $p_0 \leq 2$, it fails to produce the optimal range of exponents for $p_0 > 2$. Effectively harnessing the oscillation in the fractal analogue of the problem to obtain generalizations of [8, Theorem

3] requires new ideas and presents substantial technical challenges. We plan to return to this in a future project.

5. Our estimates, though sharp for general M and Σ, can be improved in special situations. This will be the case, for instance, when $M = \mathbb{T}^n$, the n-dimensional flat torus (which admits a stronger Weyl law), or if the submanifold Σ in a general manifold M has additional geometric properties, for example if Σ is a curve of nonvanishing geodesic curvature. This is consistent with similar results of this type for smooth submanifolds, see for example [8, Theorem 2], [4, 7, 13, 3] and the bibliography therein. We pursue this direction in greater detail in upcoming work.

6. The blow-up factor $\Phi(\lambda)$, which is super-logarithmic but sub-polynomial, is an artifact of the choices of parameters needed for the random Cantor construction, see Section 2.6. Many alternative parameter choices are possible within the framework of this construction, some of which yield logarithmic blow-up in lieu of $\Phi(\lambda)$, at the cost of additional technical challenges. We have opted not to pursue these improvements here. However, all estimates of this type will be accompanied by some blow-up. It is an interesting question whether there exists a member of this class of random sets for which such losses can be avoided.

7. The random measures ν_ω that we construct and their supporting sets Γ_ω have many analytic and geometric properties that are not directly exploited in the proof. In particular, these measures have optimal Fourier decay subject to the Hausdorff dimension of their support. More precisely, for almost every ω, our measures obey

$$\left|\widehat{\nu}_\omega(\xi)\right| \leq C_\xi \big(1 + |\xi|\big)^{-d(1-\varepsilon)/2}, \quad |\xi| \geq 1,$$

where C_ξ is a function that grows slower than any positive power of $|\xi|$. In other words, the sets Γ_ω in Theorem 1.3 have the same Fourier dimension as their Hausdorff dimension, i.e. they are almost surely Salem.

Fourier decay of measures have long been known to play an important role in eigenfunction restriction problems. For instance, it appears in the work of Bourgain and Rudnick [7], where the authors obtain significant improvements on the general estimates of [8] in the special case of $M = \mathbb{T}^n$, $n = 2, 3$. More generally, the study of harmonic-analytic principles (such as Fourier decay, fractal analogues of the uncertainty principle, study of oscillatory integrals and operators) in settings where standard techniques (such as integration by parts or stationary phase) are not viable have led to major developments in spectral theory, for instance in the work surrounding resonance gaps in infinite-area hyperbolic surfaces [25, 5, 6]. We explore the mapping properties of convolution operators on random Cantor measure spaces, and establish Young-type inequalities for such measures. However, our methods are not Fourier-analytic in nature. This is another point of similarity of our work with [22], where a similar random Cantor set was constructed, but whose Fourier-analytic properties were not directly relevant to the proof.

8. Restriction of eigenfunctions to fractals has appeared in a related but distinct line of inquiry that addresses spatial equidistribution of eigenfunctions restricted to subsets of manifolds, instead of norm growth. Most recently Hezari and Rivière [15] have established, in the specific setting of the flat torus $\mathbb{T}^n$, spatial equidistribution of a density one

subsequence of eigenfunctions on sets of possibly fractional Hausdorff dimension. In fact, Corollary 2.7 of [15] proves L^2-restriction estimates of this density one subsequence for a wider class of measures that the one used in this paper. This addresses a special case of a conjecture in [7] on L^2-restriction to hypersurfaces. It would be of interest to explore the issue of equidistribution for the entire sequence of toral eigenfunctions for the class of random Cantor sets considered in this paper.

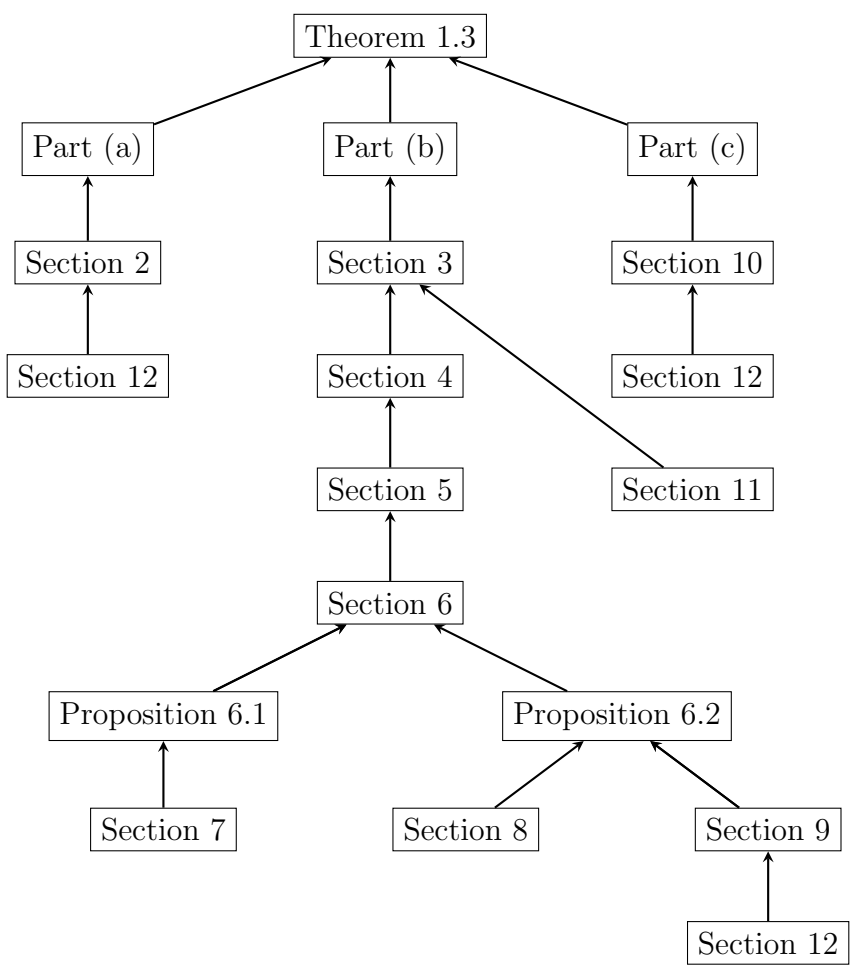
1.4. **Overview of the proof.** The broad strokes of our approach follow that of [8, Theorem 1], so we briefly review the main ideas involved here.

1. One starts with a microlocal approximation $\mathscr{T}_\lambda$ of the smoothed spectral projector $\rho(\lambda - \sqrt{-\Delta_g})$. The approximation $\mathscr{T}_\lambda$ is an oscillatory integral operator, whose phase function is essentially the distance function in the ambient Riemannian metric.
2. The TT^* method applied to $\mathscr{T}_\lambda\mathscr{T}_\lambda^*$ reduces the problem to estimating the L^p of the latter operator on the restricted set γ, which for [8, Theorem 1] was a smooth curve on M.
3. The integration kernel of $\mathscr{T}_\lambda\mathscr{T}_\lambda^*$ is itself an oscillatory integral, with a nondegenerate phase function. The method of stationary phase, applied to this oscillatory integral, yields a pointwise upper bound on the kernel, leading to a pointwise bound on the operator $\mathscr{T}_\lambda\mathscr{T}_\lambda^*$. The dominating operator is a convolution, with an explicit convolving factor.
4. The proof is then completed by invoking Young's convolution inequality for the Lebesgue measure on $\mathbb{R}$. The admissible exponents of the inequality are precisely those for which the convolving factor is integrable.

A careful analysis of [8, Theorem 1], which we carry out in Section 3, shows that steps 1, 2 and 3 above extend with minor revisions to the setting of an arbitrary measure space, with γ replaced by Γ_ω. A noteworthy point of departure is the following. Whereas the natural measure on the curve γ used in [8] is absolutely continuous with respect to the translation-invariant Lebesgue measure on $\mathbb{R}$, the measure ν_ω accompanying our Cantor set Γ_ω is no longer translation invariant. The proof thus fails critically at the last step, since Young's convolution inequality is unavailable, indeed known to be false, in general measure spaces. The main contribution of this article is in deriving an analogue of Young's inequality for the convolution kernel $\mathcal{K}_\lambda$ that appears in the pointwise upper bound in step 3, and for the special class of random Cantor measures constructed earlier in the paper (in Section 2). Specifically, this involves estimation of the quantity $\sup\big\{\|\mathcal{K}_\lambda(u-\cdot)\|_{L^p(\nu_\omega)} : u \in \Gamma_\omega\big\}$ for almost every $\omega \in \Omega$. The transition from the desired operator norm of $\mathscr{T}_\lambda\mathscr{T}_\lambda^*$ to the quantity above has been formalized in Proposition 3.3, aided in turn by a generalized Schur-type inequality proved in Section 11. A substantial portion of the article, ranging from Sections 4 to 9, is devoted to the estimation of this last quantity, through a series of successive reduction to various random sums. Critical elements of our analysis include large deviation inequalities due to Bernstein and Azuma (summarized at the end of Section 12), discretization of continuous random variables represented by the random integrals (see Section 7.2) and interaction between length scales of the fractal construction with the spectral parameter λ.

The layout of our article and logical dependencies among various sections are described in the diagram below. Section 2 is dedicated to the construction of a probability space of random

Cantor-type sets $E_\omega \subseteq [0,1]^d$, whose elements are almost surely of dimension $d(1-\varepsilon)$; see Lemma 2.4. These sets E_ω, when mapped to the d-dimensional submanifold $\Sigma \subseteq M$ via a coordinate chart, yield the desired sets $\Gamma_\omega \subseteq \Sigma$. The accompanying probability measure ν_ω is the push-forward of the natural Cantor measure μ_ω on E_ω. The proof of Theorem 1.3 (a) appears as a consequence of Lemma 2.4 in Section 2.8. Section 3 sets up the microlocal analysis background regarding the smooth spectral projector, providing in particular an explicit asymptotic expansion whose leading term becomes the main object of interest. The proof of Theorem 1.3 (b) appears here, modulo an important integration kernel estimate in Proposition 3.4 whose proof is taken up in subsequent sections. This section also contains our adaptation of Young's inequality for general measures in Proposition 3.3, with its proof relegated to Section 11. The technical work on estimating $\sup\big\{\|\mathcal{K}_\lambda(u-\cdot)\|_{L^p(\nu_\omega)} : u \in \Gamma_\omega\big\}$ begins in Section 4 where we introduce a deterministic function $\Theta(\lambda; s, \varrho; \kappa)$. This function Θ will be shown to control all the successive approximations of our integration kernel in later sections. Sections 5 through 9 encompass these approximating steps. Section 10 proves sharpness of Theorem 1.3, namely part (c), for p above the critical exponent. The key probabilistic tools and associated results that have been used repeatedly throughout the article have been collected in Section 12.



2. Preliminaries

2.1. **A general Cantor-type construction.** All the fractal subsets of $[0,1]^d$ considered in this paper are obtained using a Cantor-type iteration, whose basic features we now describe. There are two main ingredients in the construction; namely, a choice of successive scales and a selection mechanism at each scale.

Fix a nondecreasing sequence of positive integers $\{N_k : k \geq 1\}$ with

$$\delta_k^{-1} = M_k = N_1 N_2 \dots N_k. \tag{2.1}$$

Using the notation $\mathbb{Z}_m := \{1, \cdots, m\}$, we define a class of multi-indices

$$\mathbb{I}(k,d) := \big\{\mathbf{i}_k = (\bar{i}_1, \dots, \bar{i}_k);\ \bar{i}_j \in \mathbb{Z}^d_{N_j},\ 1 \leq j \leq k\big\}, \text{ and} \tag{2.2}$$

$$\mathbb{I}^* := \bigcup \big\{\mathbb{I}(k,d) : k \geq 1\big\}. \tag{2.3}$$

The interpretation of the integers N_k and the multi-indices $\mathbf{i}_k$ is the following. At step k, the unit cube $[0,1]^d$ is partitioned into subcubes of sidelength δ_k with sides parallel to the coordinate axes. These subcubes, which we term *cubes of the k-th generation*, are indexed by $\mathbf{i}_k$. Each such cube is of the form

$$\begin{aligned} Q(\mathbf{i}_k) &= \alpha(\mathbf{i}_k) + [0, \delta_k]^d, \quad \text{with} \\ \alpha(\mathbf{i}_k) &= \frac{\bar{i}_1 - \bar{1}}{N_1} + \frac{\bar{i}_2 - \bar{1}}{N_1 N_2} + \cdots + \frac{\bar{i}_k - \bar{1}}{N_1 \dots N_k}. \end{aligned} \tag{2.4}$$

Here $\bar{1} = (1, \dots, 1) \in \mathbb{R}^d$. The expression (2.4) above should be thought of as a finite "digit expansion" of $\alpha(\mathbf{i}_k)$ with respect to the base string $(N_1, N_2, \cdots)$. Every point in the unit cube has a possibly infinite digit expansion with respect to this base sequence. Further, such a digit expansion is unique, except for countably many points in the unit cube. We note that the cubes of any given generation have disjoint interiors. Further, each k-th generation cube gives rise to exactly N_{k+1}^d children, as follows:

$$Q(\mathbf{i}_k) = \bigcup \big\{Q(\mathbf{i}_k, \bar{i}_{k+1}) :\ \bar{i}_{k+1} \in \mathbb{Z}^d_{N_{k+1}}\big\}.$$

Thus any two distinct cubes $Q(\mathbf{i})$ and $Q(\mathbf{j})$ with $\mathbf{i}, \mathbf{j} \in \mathbb{I}^*$ must satisfy exactly one of the relations

$$Q(\mathbf{i}) \subsetneq Q(\mathbf{j}), \quad \text{or} \quad Q(\mathbf{j}) \subsetneq Q(\mathbf{i}), \quad \text{or} \quad \operatorname{int}(Q(\mathbf{i})) \cap \operatorname{int}(Q(\mathbf{j})) = \emptyset.$$

To specify a selection algorithm, we fix for each $k \geq 1$ an ordered set $\mathbf{Y}_k := \{Y_k(\mathbf{i}_k); \mathbf{i}_k \in \mathbb{I}(k,d)\}$ whose elements are either 0 or 1. Set $X_1(\bar{i}_1) := Y_1(\bar{i}_1)$,

$$\begin{aligned} X_k(\mathbf{i}_k) &:= X_{k-1}(\mathbf{i}_{k-1}) Y_k(\mathbf{i}_k) \text{ where } \mathbf{i}_k = (\mathbf{i}_{k-1}, \bar{i}_k), \\ \mathbf{X}_k &:= \{X_k(\mathbf{i}_k) : \mathbf{i}_k \in \mathbb{I}(k,d)\}, \quad P_k := \#\{\mathbf{i}_k : X_k(\mathbf{i}_k) = 1\}, \\ \mathcal{Q}_k &:= \{Q(\mathbf{i}_k) : \mathbf{i}_k \in \mathbb{I}(k,d), X_k(\mathbf{i}_k) = 1\}, \quad \mathcal{Q}^* := \bigcup_k \mathcal{Q}_k. \end{aligned} \tag{2.5}$$

The relevance of these definitions is the following. A total of P_k cubes of the k-th generation are chosen at step k, the marker of selection being $X_k(\mathbf{i}_k) = 1$. We call the selected ones the *basic cubes* of the k-th generation. The collection $\mathcal{Q}_k$, which is indexed by $\mathbf{i}_k$ with $X_k(\mathbf{i}_k) = 1$, specifies the cubes $Q(\mathbf{i}_k)$ that are selected. If $X_k(\mathbf{i}_k) = 0$, then so is $X_\ell(\mathbf{i}_\ell)$ for any $\ell > k$

with $\pi_k(\mathbf{i}_\ell) = \mathbf{i}_k$, by (2.5). Here π_k denotes the projection onto the first k vector coordinates in $\mathbb{R}^d$. Thus, once a cube is discarded at a given step, its descendants are eliminated from consideration for the remainder of the construction. The union of the cubes in $\mathcal{Q}_k$ therefore gives rise to a decreasing sequence of closed sets.

Given these quantities, we define the successive nested iterates E_k of the construction, and the limiting set E:

$$E_0 := [0,1]^d, \quad E_k := \bigcup\{Q(\mathbf{i}_k) : X_k(\mathbf{i}_k) = 1\} = \bigcup\{Q : Q \in \mathcal{Q}_k\}, \quad E := \bigcap_{k=1}^{\infty} E_k. \tag{2.6}$$

We always assume that $|E_k| = P_k \delta_k^d \to 0$, so that E is a Lebesgue-null set. On the other hand,

$$E \neq \emptyset \text{ if and only if } P_k \neq 0 \text{ for each } k \geq 1. \tag{2.7}$$

2.2. **A Cantor measure.** Our next task is to define a probability measure μ on E. This is a standard procedure, so we briefly sketch the details. For each $k \geq 1$, we define the function

$$\mu_k := 1_{E_k}/|E_k| = \frac{1}{P_k \delta_k^d} \sum_{\mathbf{i}_k} X_k(\mathbf{i}_k) 1_{Q(\mathbf{i}_k)}, \tag{2.8}$$

which is a probability density function that assigns a uniform mass of $1/P_k$ to each basic cube at step k. It is easy to see that the sequence $\{\mu_\ell(Q) : \ell \geq 1\}$ converges for all cubes $Q \in \mathcal{Q}^*$. Let us denote its limit by $\mu_0(Q)$, and observe that

$$\mu_0(Q) = \lim_{\ell \to \infty} \mu_\ell(Q) = \mu_k(Q) = \frac{1}{P_k} \text{ for all } Q \in \mathcal{Q}_k.$$

The set function μ_0 initially defined on $\mathcal{Q}^*$ is a pre-measure. By the Carathéodory extension theorem (see for example Proposition 1.7 in [11]), there exists a Borel probability measure μ on $[0,1]^d$ given by

$$\mu(A) := \inf\Big\{\sum_i \mu_0(U_i) : A \cap E \subseteq \bigcup U_i,\ U_i \in \mathcal{Q}^*\Big\}. \tag{2.9}$$

The measure μ coincides with μ_0 on $\mathcal{Q}^*$. In particular, $\mu_k \to \mu$ in the weak-* topology, i.e. for all $f \in C\left([0,1]^d\right)$

$$\int f\, \mu_k \to \int f\, \mu \text{ as } k \to \infty.$$

2.3. **Hausdorff dimension.** The set E defined in (2.6) obeys certain dimensionality bounds given in terms of the construction parameters.

Lemma 2.1. *Let* $\dim_{\mathbb{H}}(E)$ *denote the Hausdorff dimension of* E *constructed above. Then*

$$\dim_{\mathbb{H}}(E) \leq \liminf_{k \to \infty} \frac{\log(P_k)}{-\log(\delta_k)}, \tag{2.10}$$

$$\dim_{\mathbb{H}}(E) \geq s_0 := \liminf_{k \to \infty} \frac{\log(P_k/N_k^d)}{-\log(\delta_{k-1})}. \tag{2.11}$$

Proof. The relation (2.10) is a standard result that follows immediately from the statement of Proposition 4.1 in [11]. The proof of (2.11) is an easy adaptation of Lemma 2.1 of [22], which we include for completeness. Our main tool here is Frostman's lemma (see for example Theorem 8.8 in [24] or Section 4.1 of [11]), which says

$$\dim_{\mathbb{H}}(E) = \sup\left\{ s : \begin{array}{l} \exists \text{ a probability measure } \nu \text{ supported on } E \text{ and a constant } 0 < C < \infty \\ \text{such that } \nu(B(x;r)) \leq Cr^s \text{ for all } x \in \mathbb{R}^d \text{ and } r > 0. \end{array} \right\}.$$

Since any ball of radius r can be covered by an axes-parallel cube of sidelength $2r$, the claim (2.11) would follow if we prove the following: for μ defined as in Section 2.2 and every $s < s_0$, there is a constant $C_s > 0$ for which the estimate

$$\mu(J) \leq C_s r^s \tag{2.12}$$

holds for all axes-parallel cubes J of sidelength r. To this end, fix a small number $r > 0$ and let $k = k(J)$ denote the unique index such that $\delta_{k+1} \leq r < \delta_k$. The number of basic cubes in $\mathcal{Q}_{k+1}$ that can intersect J is either: (i) at most $2^d N_{k+1}^d$ as J may intersect at most 2^d adjacent cubes in $\mathcal{Q}_k$, or (ii) the natural upper bound of $|J|/\delta_{k+1}^d$, since the cubes in $\mathcal{Q}_{k+1}$ have disjoint interiors. From the definition (2.9) of μ, we see that

$$\begin{aligned} \mu(J) &\leq P_{k+1}^{-1} \min\left[2^d N_{k+1}^d, \frac{|J|}{\delta_{k+1}^d}\right] \\ &\leq P_{k+1}^{-1} \left(2^d N_{k+1}^d\right)^{1-\theta} \left(\frac{|J|}{\delta_{k+1}^d}\right)^{\theta} \\ &\leq 2^d \frac{N_{k+1}^d}{P_{k+1}\delta_k^{d\theta}} r^{d\theta}. \end{aligned}$$

Here $0 \leq \theta \leq 1$ is a constant to be determined shortly. Setting $s = d\theta$, we observe that (2.12) is met provided $N_{k+1}^d / P_{k+1}\delta_k^s$ is uniformly bounded for all sufficiently large k. This happens precisely when $s < s_0$, completing the proof. □

2.4. **Random Cantor sets.** We now delve into the probabilistic construction which generates our desired Cantor-type sets. The basic procedure is as in Section 2.1, with the crucial additional point that the sequence $\mathbf{X}_k$ is now randomized. Recall the definitions of M_k, N_k and $\mathbb{I}(k,d)$ from (2.2) and the discussion preceding it.

The underlying measure space under consideration is

$$\Omega = \prod_{k=1}^{\infty} \Omega_k, \quad \text{where} \quad \Omega_k = \prod_{\mathbf{i}_k \in \mathbb{I}(k,d)} \Xi_k(\mathbf{i}_k), \quad \text{with} \quad \Xi_k(\mathbf{i}_k) = \{0,1\}.$$

Thus Ω_k consists of all binary strings of length M_k^d. We denote a "random" string in Ω_k by $\mathbf{Y}_k = \{Y_k(\mathbf{i}_k) : \mathbf{i}_k \in \mathbb{I}(k,d)\}$. The set Ω is the collection of all infinite binary strings $\omega = (\mathbf{Y}_1, \mathbf{Y}_2, \cdots)$, with $\mathbf{Y}_k \in \Omega_k$. As described in Section 2.1, every $\omega \in \Omega$ generates a Cantor-type set $E(\omega)$.

We now assign a probability measure to Ω. For a collection of small positive numbers ε_k to be specified, set

$$p_k = N_k^{-d\varepsilon_k} \tag{2.13}$$

which will serve as the "selection probability" at step k. For each $k \geq 1$ and $\mathbf{i}_k \in \mathbb{I}(k,d)$, the two-point set $\Xi_k(\mathbf{i}_k) = \{0,1\}$ is endowed with the probability measure λ_k, where

$$\begin{aligned}\lambda_k(\{1\}) &= \text{ probability of the event } \{Y_k(\mathbf{i}_k) = 1\} = p_k, \quad \text{ and}\\ \lambda_k(\{0\}) &= \text{ probability of the event } \{Y_k(\mathbf{i}_k) = 0\} = 1 - p_k.\end{aligned}$$

The space Ω_k is equipped with a product probability measure that is a finite $\#(\mathbb{I}(k,d))$-fold Cartesian product of the measures λ_k, with one copy of λ_k for each $\mathbf{i}_k \in \mathbb{I}(k,d)$. In other words,

$$\mathbb{P}_k = \prod_{\mathbf{i}_k \in \mathbb{I}(k,d)} \lambda_k.$$

This means that for every binary string $\eta = (\eta_1, \cdots, \eta_{M_k^d}) \in \Omega_k$,

$$\begin{aligned}\mathbb{P}_k(\{\eta\}) &= \text{probability of } \{\mathbf{Y}_k = \eta\} = p_k^{|\eta|}(1-p_k)^{M_k^d - |\eta|} \text{ where}\\ |\eta| &= \eta_1 + \cdots + \eta_{M_k^d} = \text{number of 1-s in the string } \eta.\end{aligned}$$

Finally, the measure $\mathbb{P}$ on Ω is the product probability measure of the measures $\mathbb{P}_k$. In summary, the random vectors $\{\mathbf{Y}_k : k \geq 1\}$ across different scales are independent. The scalar entries $\{Y_k(\mathbf{i}_k) : \mathbf{i}_k \in \mathbb{I}(k,d)\}$ within a single scale $\mathbf{Y}_k$ are independent as well; in addition, they are identically distributed as Bernoulli random variables with success probability p_k. On the other hand, it is important to note that the random variables $\{\mathbf{X}_k : k \geq 1\}$ defined as in (2.5) are *not* independent.

Before proceeding further, we need to ensure that the limiting sets $E = E(\omega)$ obtained in this manner are nonempty, with nonzero probability.

Lemma 2.2. *Assume that the construction parameters N_k and ε_k defining* (2.13) *are chosen so that*

$$\sum_{k=1}^{\infty} (1-p_k)^{N_k^d} < 1. \tag{2.14}$$

Then for the construction described above, $\mathbb{P}(E$ is nonempty$) > 0$.

Proof. It suffices to show that the probability of the complementary event, namely when E is empty, is bounded from above by the left hand side of (2.14). Accordingly, we express this event as a disjoint union:

$$\begin{aligned}\{E \text{ is empty}\} &= \bigcup_{k=1}^{\infty} \big\{\exists k \text{ such that } E_k \text{ is empty}\big\}\\ &= \dot{\bigcup} \big\{P_k = 0 \text{ but } P_{k-1} > 0\big\}\\ &\subseteq \bigcup_{k=1}^{\infty} \big\{\exists\, \mathbf{i} \in \mathbb{I}(k-1,d) \text{ such that } Y_k(\mathbf{i}_k) = 0 \;\forall\, \mathbf{i}_k = (\mathbf{i}, i_k) \in \mathbb{I}(k,d)\big\}.\end{aligned}$$

The k-th event in the last union of sets can happen only if there are at least N_k^d independent Bernoulli random variables $Y_k(\mathbf{i}_k)$ at the k-th stage that vanish. Since each $Y_k(\mathbf{i}_k)$ assumes the value 1 with probability p_k, the probability of this k-th event is at most $(1-p_k)^{N_k^d}$, completing the proof. □

In view of Lemma 2.2, we can define the conditional probability measure $\mathbb{P}^*$ as follows: for any measurable set A,

$$\mathbb{P}^*(A) := \frac{\mathbb{P}(A \cap \{E \neq \emptyset\})}{\mathbb{P}(E \neq \emptyset)}. \tag{2.15}$$

Most of the probabilistic statements made in this paper will be with respect to $\mathbb{P}^*$.

2.5. **Quantitative estimates of P_k.** We pause a moment to record some bounds on P_k that will play a crucial role in the sequel. Set

$$\overline{P}_k := N_k^{d(1-\varepsilon_k)} P_{k-1}, \qquad R_k := \prod_{j=1}^{k} N_j^{d(1-\varepsilon_j)}. \tag{2.16}$$

Note that while P_k is a random variable given by $\mathbf{X}_k$, the random variable $\overline{P}_k$ depends only on $\mathbf{X}_{k-1}$. The quantity R_k on the other hand is purely deterministic.

Lemma 2.3. *Assume that the construction parameters N_k and ε_k obey* (2.14) *and*

$$\sum_{k=1}^{\infty} \log k \; N_k^{-d(1-\varepsilon_k)/2} < \infty. \tag{2.17}$$

Then for $\mathbb{P}^*$-almost every $\omega \in \Omega$, there exist constants $C_1, C_2 \geq 1$ depending on ω such that for every $k \geq 1$, the following two estimates hold:

$$|P_k - \overline{P}_k| \leq C_1 \sqrt{\log(k+1)} \max\big(\overline{P}_k, \log(k+1)\big)^{\frac{1}{2}}, \tag{2.18}$$

$$C_2^{-1} R_k \leq P_k \leq C_2 R_k. \tag{2.19}$$

Proof. Estimates of this type are consequences of large deviation inequalities ubiquitous in the probabilistic literature, and have also appeared in previous work on random construction of sets, see for instance [21, 22]. For completeness, we include the proof of (2.18) and (2.19) in the Section 12. □

2.6. **Choice of construction parameters.** So far, we have not specified values of N_k and ε_k that are used in the random construction of our Cantor sets. We do so now. Even though a vast majority of our results will continue to hold for very general choices of large N_k and small ε_k, we set down two specific choices of (N_k, ε_k)-pairs that will be used as reference points for the rest of the analysis. They are

$$N_k := N^k, \quad \varepsilon_k = \frac{\gamma}{k}, \text{ and} \tag{2.20}$$

$$N_k := N^k, \quad \varepsilon_k = \varepsilon. \tag{2.21}$$

In (2.20), $N \geq 1$ is a fixed large integer and $0 < \gamma < 1$ is a small constant such that N^γ is large. In (2.21), N is a fixed large integer and $0 < \varepsilon < 1$ is an arbitrary constant independent of N. We leave the reader to verify that both (2.14) and (2.17) hold for these choices of N_k and ε_k, so that $\mathbb{P}^*$ is well-defined according to (2.15).

2.7. **Almost sure Hausdorff dimension.** The relevance of the choices of ε_k in (2.20) and (2.21) is clarified in the following lemma.

Lemma 2.4. *For the random construction described in Section 2.4, and for $\mathbb{P}^*$-almost every $\omega \in \Omega$, the corresponding set $E = E(\omega)$ obeys the dimensional bound*

$$dim_{\mathbb{H}}(E) = \begin{cases} d & \text{for } N_k \text{ and } \varepsilon_k \text{ as in (2.20)}, \\ d(1-\varepsilon) & \text{for } N_k \text{ and } \varepsilon_k \text{ as in (2.21)}. \end{cases} \tag{2.22}$$

Proof. Note that $M_k = \delta_k^{-1} = N^{k(k+1)/2}$ for both choices (2.20) and (2.21). In view of (2.16) and (2.19), the quantity P_k is $\mathbb{P}^*$-almost surely bounded above and below by constant multiples of

$$R_k = \prod_{j=1}^{k} N_j^{d(1-\varepsilon_j)} = \left\{ \begin{array}{ll} \displaystyle\prod_{j=1}^{k} N^{dj(1-\gamma/j)} = N^{dk(k-1)/2} & \text{for (2.20)}, \\ \displaystyle\prod_{j=1}^{k} N^{dj(1-\varepsilon)} = N^{d(1-\varepsilon)k(k+1)/2} & \text{for (2.21)}. \end{array} \right\} \tag{2.23}$$

With these choices of δ_k and R_k, it is now easy to check that

$$\liminf_{k\to\infty} \frac{\log(P_k/N_k^d)}{-\log(\delta_{k-1})} = \liminf_{k\to\infty} \frac{\log(P_k)}{-\log(\delta_k)} = \lim_{k\to\infty} \frac{\log(R_k)}{-\log(\delta_k)},$$

and that the value of the limit is precisely the quantity in the right hand side of (2.22) in the two cases. The desired conclusion now follows from Lemma 2.1. □

Remark: Lemma 2.4 says that $d\varepsilon$ should be viewed as a marker of "codimension" of the set E, which therefore is independent of γ (for (2.20)) and N (for both (2.20) and (2.21)). We therefore fix N and γ as absolute constants that will not change in the sequel; for instance, $N = 10^6$ and $\gamma = 1/3$ will suffice. The quantity ε will vary. For notational consistency, we will henceforth set $\varepsilon = 0$ for the case given by (2.20).

2.8. **Fractal subsets of manifolds.** In this subsection, we describe how to transfer our Cantor set constructions from a unit cube to the setting of a Riemannian (sub)manifold. Given a compact n-dimensional Riemannian manifold M, let $\Sigma \subset M$ be a smooth embedded (sub)manifold of dimension $1 \le d \le n$, equipped with the restricted Riemannian metric naturally endowed by g. Let (U, φ) be a local coordinate chart on Σ, where $U \subseteq \mathbb{R}^d$ is an open set containing $[0,1]^d$ and $\varphi : U \to \varphi(U) \hookrightarrow \Sigma$ is a smooth embedding. For a random Cantor set $E = E(\omega)$ with $\omega \in \Omega$ constructed as in Section 2.4, we define the corresponding random set Γ_ω in Σ by setting

$$\Gamma_\omega := \varphi(E(\omega)), \quad \text{provided} \quad E(\omega) \neq \emptyset.$$

2.8.1. *Proof of Theorem 1.3* (a).

Proof. Since φ is a diffeomorphism, it is bi-Lipschitz, and hence preserves Hausdorff dimension [11, Corollary 2.4]. Thus $\dim_{\mathbb{H}}(\Gamma_\omega) = \dim_{\mathbb{H}}(E(\omega))$. The result of Theorem 1.3 (a) now follows from Lemma 2.4. □

2.8.2. *A measure on* Γ_ω. If $\mu = \mu_\omega$ is the probability measure on $E = E(\omega)$ defined as in Section 2.2, we define the corresponding measure $\nu = \nu_\omega$ on $\Gamma = \Gamma_\omega$ as follows: for $f \in C^\infty(M)$,

$$\int_\Gamma f\, d\nu := \int_E f \circ \varphi(u) \sqrt{\det g(u)}\, d\mu(u). \tag{2.24}$$

Since $\det(g)$ is a positive smooth function that is bounded above and away from zero on U, the weight factor $\sqrt{\det(g)}$ in (2.24) is a benign one, in the sense that it can be replaced by constants in many operator-theoretic estimates. Let dV_g denote the canonical volume measure on M bestowed by g. We will study the restriction of eigenfunctions normalized in $L^2(M) = L^2(M, dV_g)$ to the Cantor set $\Gamma = \Gamma_\omega$ equipped with the measure $\nu = \nu_\omega$.

3. Microlocal preliminaries

Our study of eigenfunction restriction estimates is based on an explicit integral representation of an underlying operator $\mathscr{T}_\lambda$, called the smoothed spectral projector, in local coordinates. The formulation of the spectral projector $\mathscr{T}_\lambda$ as a Fourier integral operator is well-known and ubiquitous in the literature (see [34] and the references therein). This is stated in Theorem 3.1 below for completeness. After restricting $\mathscr{T}_\lambda$ to a random Cantor set Γ as specified in Section 2.8, we arrive at

$$\mathcal{T}_\lambda := \mathscr{T}_\lambda\Big|_\Gamma, \tag{3.1}$$

which is the main operator of interest. A microlocal analysis of the integral kernel of the resulting symmetrized operator $\mathcal{T}_\lambda \mathcal{T}_\lambda^*$ is a key tool in determining Lebesgue mapping properties of $\mathcal{T}_\lambda$. We will relate this kernel with that of the unrestricted normal operator $\mathscr{T}_\lambda \mathscr{T}_\lambda^*$. In Theorem 3.2, we recall a standard asymptotic expansion of the kernel of the latter operator. This section is devoted to a recollection of classical facts, stated largely without proof but with suitable references. Collectively, they provide a roadmap leading up to the proof of Theorem 1.3 in Section 3.5, modulo a probabilistic kernel estimate that has been specified in Proposition 3.4 and will be proved in later sections.

3.1. **The smoothed spectral projector.** Let us consider a Riemannian uniformly normal neighbourhood of a base point $x_0 \in M$. Let $\kappa > 0$ denote the radius of a geodesic ball centred at x_0 contained in this neighbourhood. Without loss of generality and due to our normal coordinate system, we can write $x_0 = 0$, the origin in $\mathbb{R}^n = T_{x_0}M$. We set $V = \{y \in \mathbb{R}^n : |y| \le \kappa\}$ and fix a neighbourhood W of x_0 contained in V.

Let us consider the first-order pseudo-differential operator $\sqrt{-\Delta}$ given by the spectral theorem. Consequently we have

$$\sqrt{-\Delta} = \sum_{j=0}^\infty \lambda_j \mathscr{P}_j, \qquad I = \sum_{j=0}^\infty \mathscr{P}_j,$$

where $\mathscr{P}_j$ is the projection operator onto the finite-dimensional eigenspace corresponding to the eigenvalue λ_j. Furthermore

$$e^{it\sqrt{-\Delta}} = \sum_{j=0}^\infty e^{it\lambda_j} \mathscr{P}_j.$$

Let us now fix a function $\chi \in \mathcal{S}(\mathbb{R}^n)$ with $\chi(0) = 1$ and $\mathrm{supp}(\widehat{\chi}) \subseteq (\kappa/2, \kappa)$. This leads to the smooth projection operators

$$\mathscr{T}_\lambda := \chi(\sqrt{-\Delta} - \lambda) = \sum_j \chi(\lambda_j - \lambda)\mathscr{P}_j. \tag{3.2}$$

We observe that

$$\chi(\sqrt{-\Delta} - \lambda)\varphi_\lambda = \varphi_\lambda \text{ for all } \lambda = \lambda_j. \tag{3.3}$$

Furthermore, a formal operator calculus made rigorous by the spectral theorem shows that

$$\begin{aligned} \mathscr{T}_\lambda := \chi(-\sqrt{-\Delta} - \lambda) &= \sum_j \chi(\lambda_j - \lambda)\mathscr{P}_j \\ &= \sum_j \left[\frac{1}{2\pi}\int e^{it(\lambda_j - \lambda)}\hat{\chi}(t)\,dt\right] \circ \mathscr{P}_j \\ &= \frac{1}{2\pi}\int e^{-it\lambda}e^{it\sqrt{-\Delta}}\hat{\chi}(t)\,dt. \end{aligned} \tag{3.4}$$

For κ small enough, it is a classical result that $e^{it\sqrt{-\Delta}}$ can be represented as a Fourier integral operator in local coordinates, see for instance [16]. A stationary phase argument then leads to the following well-known theorem:

Theorem 3.1. [34, Theorem 4.1.2, Lemma 5.1.3] *Given the setup described above, there exist distinct positive constants $c_i \in (0,1)$, $0 \le i \le 4$, and a smooth function $a : W \times V \times \mathbb{R}^+_\lambda \to \mathbb{C}$ with support in the set*

$$S := \big\{(x,y) \in W \times V : |x| \le c_0\kappa < c_1\kappa \le |y| \le c_2\kappa < \kappa\big\} \tag{3.5}$$

obeying the following properties for $\lambda \ge 1$:

(a) The function $a_\lambda(x,y) := a(x,y;\lambda)$ does not vanish for $(x,y) \in S$ with $|x| \le c_0\kappa$ and $d_g(x,y) \in [c_3\kappa, c_4\kappa]$.

(b) The spatial derivatives of a_λ are uniformly bounded, i.e., for every multi-index α, there exists a constant $C_\alpha > 0$ such that $|\partial^\alpha_{x,y} a_\lambda(x,y)| \le C_\alpha$.

(c) The function a_λ appears in the representation of the integral kernel for the smoothed spectral projector $\mathscr{T}_\lambda$ defined as in (3.4). Specifically, for all $x \in W$ and all $f \in L^2(V)$

$$\mathscr{T}_\lambda(f)(x) = \lambda^{\frac{n-1}{2}} \int_{y \in V} e^{-i\lambda d_g(x,y)} a_\lambda(x,y) f(y)\,dy + \mathscr{R}_\lambda(f). \tag{3.6}$$

Here $\mathscr{R}_\lambda$ is a smoothing operator in the sense that $||\mathscr{R}_\lambda||_{L^2(V)\to L^q(W)} \le C_N\lambda^{-N}$ for all $N \ge 1$ and all $2 \le q \le \infty$.

The infinitely smoothing property of $\mathscr{R}_\lambda$ ensures that for any probability measure μ supported on a set $\Gamma \subseteq M$, we have the following estimate

$$||\mathscr{R}_\lambda f||_{L^2(\mu)} \le ||\mathscr{R}_\lambda f||_{L^\infty(\mu)} \le ||\mathscr{R}_\lambda f||_{L^\infty(W)} \le C_N\lambda^{-N}||f||_{L^2(M)}$$

for all smooth functions $f \in L^2(V)$ and all $N \ge 1$. Thus $\mathscr{R}_\lambda$ does not contribute any significant power of λ in eigenfunction restriction estimates, and we ignore it in the sequel. By a slight abuse of notation, we will rename as $\mathscr{T}_\lambda$ the leading term in (3.6).

3.2. **Restriction of the smoothed spectral projector to submanifolds.** Let us recall from Sections 2.8 and 2.2 the construction of a random Cantor set Γ on a (sub)manifold $\Sigma \subseteq M$, endowed with its natural measure ν. Equipped with the representation (3.6), we now embark on the study of the smooth spectral projection operator $\mathscr{T}_\lambda$ *restricted* to Γ, as defined by (3.1). By duality, $\mathcal{T}_\lambda$ maps $L^2(M, dV)$ boundedly to $L^p(\Gamma, \nu)$ if and only if $\mathcal{T}_\lambda \mathcal{T}_\lambda^*$ maps $L^{p'}(\Gamma, \nu)$ to $L^p(\Gamma, \nu)$. We are thus led to examine the latter normal operator.

Let $\mathscr{K}_\lambda(x, x')$ denote the Schwarz kernel of the operator $\mathscr{T}_\lambda \mathscr{T}_\lambda^*$, where $\mathscr{T}_\lambda$ is the operator given by the leading term in (3.6). If (U, φ) is a local coordinate chart on Σ as specified in Section 2.8, it follows that

$$\mathcal{T}_\lambda \mathcal{T}_\lambda^*(f \circ \varphi)(u) = \int \mathscr{K}_\lambda(\varphi(u), \varphi(v)) f \circ \varphi(v) \sqrt{\det(g(v))}\, d\mu(v), \qquad u \in U,\ \varphi(u) \in \Sigma. \tag{3.7}$$

In other words, the Schwarz kernel of $\mathcal{T}_\lambda \mathcal{T}_\lambda^*$ is the restriction of $\mathscr{K}_\lambda$ to $\Gamma \times \Gamma$. The first step in establishing Lebesgue boundedness of the normal operator $\mathcal{T}_\lambda \mathcal{T}_\lambda^*$ is therefore studying the Schwarz kernel $\mathscr{K}_\lambda$. Our next result gives an asymptotic expansion of this kernel.

Theorem 3.2. [8, Lemma 6.1] *For $\mathscr{K}_\lambda(x, x')$ be as in the preceding paragraph, the following conclusions hold:*

(a) There exist constants $\kappa < 1 < C$ and a sequence of real-valued symbols $(a_m^\pm, b_m) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R})$ such that for $|x - x'| \geq C\lambda^{-1}$ and any $N \in \mathbb{N}$, the following expansion holds:

$$\lambda^{1-n} \mathscr{K}_\lambda(x, x') = \sum_{\pm} \sum_{m=0}^{N-1} \frac{e^{\pm i\lambda d_g(x,x')}}{(\lambda|x - x'|)^{\frac{n-1}{2}+m}} a_m^\pm(x, x', \lambda) + b_N(x, x', \lambda). \tag{3.8}$$

Each of the symbols $a_m^\pm$ has support in $W \times V$ of size $\mathcal{O}(\kappa)$ (independent of λ) and is uniformly bounded in λ. The remainder b_N obeys the estimate

$$|b_N(x, x', \lambda)| \leq C_N\, (\lambda|x - x'|)^{\frac{n-1}{2} - N}.$$

(b) In particular, if $\{x = \varphi(u) : u \in U\}$ is a local parameterization of a d-dimensional smooth embedded (sub)manifold $\Sigma \subseteq M$, then there exists a constant $C > 1$ such that

$$\left|\mathscr{K}_\lambda(x, x')\right| \leq C\lambda^{n-1} \langle \lambda(u - v)\rangle^{-\frac{n-1}{2}}, \tag{3.9}$$

for all $x = \varphi(u)$, $x' = \varphi(v) \in \Sigma$, $u, v \in U$. Here $\langle \cdot \rangle$ denotes the Japanese bracket given by $\langle u \rangle := (1 + |u|^2)^{1/2}$.

3.3. **Reduction to a generalized Young-type inequality.** The kernel estimate (3.9) allows us to bound $\mathcal{T}_\lambda \mathcal{T}_\lambda^*$ pointwise by a convolution operator.

Proposition 3.3. *Given a set $\Gamma \subseteq \Sigma$, let $\mathcal{T}_\lambda$ denote the restriction to Γ of the smooth spectral projection operator, as defined by (3.1). Suppose that Γ is parameterized by $E \subseteq [0, 1]^d$ via a coordinate chart φ, i.e., $\Gamma = \varphi(E)$.*

(a) Then for all non-negative f and all $u \in E$, we have the pointwise inequality

$$\left|\mathcal{T}_\lambda \mathcal{T}_\lambda^*(f \circ \varphi)(u)\right| \leq C\lambda^{n-1} \big[\mathcal{L}_\lambda(f \circ \varphi)\big](u), \quad \textit{where} \tag{3.10}$$

$$\mathcal{L}_\lambda(f \circ \varphi)(u) := \int \mathcal{K}_\lambda(u - v)(f \circ \varphi)(v)\, d\mu(v), \ \textit{and}$$

$$\mathcal{K}_\lambda(u) := \langle \lambda u \rangle^{-\frac{n-1}{2}}.$$

(b) For $p \geq 2$, the operator $\mathcal{T}_\lambda$ is bounded as a linear operator from $L^2(M, dV)$ to $L^p(\Gamma, \nu)$ provided

$$a_p^{\frac{p}{2}} := \sup_{u \in E} \int \left|\mathcal{K}_\lambda(u-v)\right|^{\frac{p}{2}} d\mu(v) < \infty. \tag{3.11}$$

The operator norm of $\mathcal{T}_\lambda$ is, in this case, bounded above by $\sqrt{\lambda^{n-1} a_p}$.

Proof. The inequality (3.10) follows from (3.7) combined with (3.9), once we recall from Section 2.8 that the weight factor $\sqrt{\det(g)}$ is uniformly bounded above and below by positive constants. Part (b) is a consequence of a generalized Young's inequality, stated and proved in Proposition 11.1 in an appendix below (Section 11). We have used this proposition with T replaced by $\mathcal{L}_\lambda$, $r = p$, $q = p'$ and $s = p/2$. In view of the symmetry and the translation-invariance of the kernel $\mathcal{K}_\lambda$, both the quantities A_s and B_s in (11.3) equal a_p in this context. If a_p is finite, the conclusion of Proposition 11.1 asserts that $\lambda^{n-1}\mathcal{L}_\lambda$ is bounded as a linear operator from $L^{p'} \to L^p$ with norm at most $\lambda^{n-1} a_p$. In view of (3.10), $\mathcal{T}_\lambda \mathcal{T}_\lambda^*$ has the same property. By duality, $\mathcal{T}_\lambda$ maps $L^2(M, dV_g)$ to $L^p(\Gamma, \nu)$ with norm bounded by the square root of $\lambda^{n-1} a_p$. □

3.4. **An integration kernel estimate.** In view of Proposition 3.3, the problem of Lebesgue boundedness of $\mathcal{T}_\lambda$ reduces to an estimation of the quantity a_p in (3.11). For a random set Γ and its associated measure ν as described in Section 2, the integral representing a_p is random, so we aim to prove a quantitative estimate for it that holds almost surely. The following proposition, which makes this precise, is the main step towards Theorem 1.3.

For $p \in (0, \infty]$, set

$$p_0 := \frac{4d(1-\varepsilon)}{n-1}, \tag{3.12}$$

$$\alpha_p = \alpha(p, n, d, \varepsilon) := \left\{ \begin{array}{l} \dfrac{p(n-1)}{4} \text{ if } 0 < p \leq p_0, \\ d(1-\varepsilon) \text{ if } p \geq p_0. \end{array} \right\} \tag{3.13}$$

Given $\mu = \mu_\omega$, the natural measure on the random Cantor set $E = E_\omega$ described in Section 2.4, we define the random function

$$\mathfrak{A}(u, \lambda; p) = \mathfrak{A}_\omega(u, \lambda; p) := \int \langle \lambda(u-v) \rangle^{-\frac{p(n-1)}{4}} d\mu(v). \tag{3.14}$$

The most important technical component of this article is the following proposition.

Proposition 3.4. *Fix any $p \in [1, \infty)$. For $\mathbb{P}^*$-almost every ω, there exists a constant $C > 0$ depending only on $\omega, p, n, d, \varepsilon$ such that for all $\lambda \geq 1$, the following estimate holds:*

$$\sup_{u \in [0,1]^d} |\mathfrak{A}(u, \lambda; p)| \leq C \Phi(\lambda) \lambda^{-\alpha_p}. \tag{3.15}$$

Here $\mathfrak{A}$, α_p and Φ are as in (3.14), (3.13) *and* (1.9) *respectively.*

3.5. **Proof of Theorem 1.3 (b), assuming Proposition 3.4.**

Proof. Let us recall the relation (3.3), which in particular implies $\mathscr{T}_\lambda \varphi_\lambda = \varphi_\lambda$ for every eigenfunction φ_λ. Thus (1.7) follows from the stronger statement

$$||\mathcal{T}_\lambda||_{L^2(M)\to L^p(\Gamma,\nu)} \leq C\Phi(\lambda)(1+\lambda)^{\delta_p},$$

with δ_p as given by Theorem 1.3 (b). We set about proving this.

For $p_0 > 2$, we will apply Proposition 3.4 twice, for the values $p = 2$ and $p = p_0$ respectively. The proposition then gives that except for a $\mathbb{P}^*$-null set of ω, the random function $\mathfrak{A}_\omega(\cdot, \lambda; p)$ obeys the estimate (3.15) for these two values of p. In the notation of Proposition 3.3 part(b), this means that for $p = 2$ and $p = p_0$, the operator $\mathcal{T}_\lambda$ maps $L^2(M, dV_g)$ to $L^p(\Gamma, d\nu)$, with operator norm bounded above by $\sqrt{\lambda^{n-1} a_p}$, where

$$a_p \leq \left(C\Phi(\lambda)\lambda^{-\alpha_p}\right)^{\frac{2}{p}} = \lambda^{-\frac{n-1}{2}}\Phi(\lambda)$$

in each case. The conclusion of Theorem 1.3 for $p_0 > 2$ with $2 \leq p \leq p_0$ now follows from Hölder's inequality, by interpolating the linear operator $\mathcal{T}_\lambda$ between these two values of p. For $p_0 > 2$ and $p \in [p_0, \infty]$, we interpolate $\mathcal{T}_\lambda$ between $p = p_0$ and the trivial bound at $p = \infty$, namely

$$||\mathcal{T}_\lambda f||_\infty \leq C\lambda^{\frac{n-1}{2}}||f||_2.$$

The last inequality follows from the Weyl law, and is also an easy consequence of (3.6). For $p_0 \leq 2$, it is only necessary to interpolate once, between the endpoints $p = 2$ and $p = \infty$, completing the proof of the theorem. □

4. Approximation of the integration kernel: Proof of Proposition 3.4

4.1. **Notation.** We have seen in Section 3.5 that the proof of Theorem 1.3 is predicated on Proposition 3.4. We are thus tasked with proving the almost sure estimate (3.15). In this section we take a step in this direction, by recording a probabilistic statement concerning certain approximations of $\mathfrak{A}$ that ultimately lead to (3.15). This statement is contained in Proposition 4.2 which is the main result of this section, and the proof of Proposition 3.4 is completed using it. The proof of Proposition 4.2 will be presented in Section 5.

We begin by setting up some preparatory notation. For ease of exposition, it is convenient to define the following deterministic function Θ, which will dominate various quantities that we will estimate. For fixed positive constants s and ϱ, we define

$$\Theta(\lambda; s, \varrho; \kappa) := \begin{cases} \kappa_\lambda \lambda^{-\varrho} & \text{if } s \geq \varrho, \\ \lambda^{-s} & \text{if } s < \varrho. \end{cases} \tag{4.1}$$

Here $\kappa = \kappa_\lambda$ for $\lambda \geq 1$ is a monotone nondecreasing function of λ that grows to infinity slower than any power of λ. Our analysis will show that for functions Θ relevant to Proposition 3.4, the exact functional form of κ_λ is tied to the choice of parameters (2.20) and (2.21) in the Cantor construction. As long as κ_λ grows slower than any power of λ, it does not affect the exponent α_p of λ in Proposition 3.4; but for specificity we will keep track of it nonetheless.

We ask the reader to verify that the bound on $\mathfrak{A}$ specified on the right hand side of the inequality (3.15) is in fact $\Theta(\lambda; s, \rho; \kappa)$, with

$$s = \frac{p(n-1)}{4}, \quad \varrho = d(1-\varepsilon), \quad \kappa_\lambda = \Phi(\lambda). \tag{4.2}$$

We start by recording an easy integral estimate in terms of Θ that will be used extensively in the sequel.

Lemma 4.1. *Given numbers $\vartheta, \rho > 0$, there exists a constant $C = C(\vartheta, \rho)$ such that for all $\tau > 0$,*

$$\int_0^\tau t^{\vartheta-1}\langle t\rangle^{-\rho}\,dt \le C\tau^\vartheta \times \begin{cases} 1 & \text{if } 0 < \tau \le 1, \\ \Theta(\tau; \rho, \vartheta; \log(1+\tau)) & \text{if } \tau > 1. \end{cases} \tag{4.3}$$

Proof. A direct computation shows that

$$\int_0^\tau t^{\vartheta-1}\langle t\rangle^{-\rho}\,dt \le C \times \begin{cases} \tau^\vartheta & \text{if } \tau < 1, \\ 1 & \text{if } \tau \ge 1 \text{ and } \vartheta < \rho, \\ \log(1+\tau) & \text{if } \tau > 1 \text{ and } \vartheta = \rho, \\ \tau^{\vartheta-\rho} & \text{if } \tau > 1 \text{ and } \vartheta > \rho. \end{cases}$$

It is now straightforward to verify that the last expression above is bounded by the right hand side of (4.3). □

4.2. **Approximating $\mathfrak{A}$ using $\mathscr{A}_k$ and $\mathscr{B}_k$.** The main quantity of interest $\mathfrak{A}$ will in turn be approximated using absolutely continuous approximations of the measure μ. More precisely, for any fixed positive constant s, set

$$\mathscr{A}_k(u, \lambda; s) = \mathscr{A}_{k,\omega}(u, \lambda; s) := \int \langle \lambda(u-v)\rangle^{-s}\,d\mu_k(v), \quad \text{and} \tag{4.4}$$

$$\mathscr{A}(u, \lambda; s) = \mathscr{A}_{\omega}(u, \lambda; s) := \int \langle \lambda(u-v)\rangle^{-s}\,d\mu(v). \tag{4.5}$$

Thus

$$\mathfrak{A}(u, \lambda; p) = \mathscr{A}(u, \lambda; p(n-1)/4), \quad \text{with} \quad s = p(n-1)/4. \tag{4.6}$$

The measure μ_k used in the definitions above is the normalized Lebesgue measure on the k-th Cantor iterate E_k, as defined in (2.8) of Section 2.2. Substituting (2.8) into the expression for $\mathscr{A}_k$ in (4.4), we obtain

$$\mathscr{A}_k = \mathscr{A}_k(u, \lambda; s) = (P_k\delta_k^d)^{-1}\mathscr{B}_k(u, \lambda; s), \text{ where} \tag{4.7}$$

$$\mathscr{B}_k = \mathscr{B}_k(u, \lambda; s) := \sum_{\mathbf{i}_k} X_k(\mathbf{i}_k) w_k(\mathbf{i}_k), \text{ and} \tag{4.8}$$

$$w_k(\mathbf{i}_k) = w_k(\mathbf{i}_k; u, \lambda, s) := \int_{Q(\mathbf{i}_k)} \langle \lambda(u-v)\rangle^{-s}\,dv. \tag{4.9}$$

The relevance of the quantities above in the estimation of $\mathfrak{A}$ is the following. The quantity $\mathscr{A}_k$ (respectively $\mathscr{A}$) represents the action of the measure μ_k (respectively μ) on the continuous

function $v \mapsto \langle \lambda(u-v)\rangle^{-s}$. Since μ_k converges to μ in the weak-$*$ topology, we have that

$$\mathscr{A}_k(u,\lambda;s) \to \mathscr{A}(u,\lambda;s) \quad \text{as } k\to\infty, \text{ for every } u\in[0,1]^d \text{ and } \lambda>0. \tag{4.10}$$

Thus any upper bound on $\mathscr{A}$ or $\mathfrak{A}$ would follow from a similar estimate on $\mathscr{A}_k$, for all sufficiently large k. The precise estimate is formalized in the proposition below.

Proposition 4.2. *For $\mathbb{P}^*$-almost every ω, there exists a constant C_1 depending only on $\omega, p, n, d, \varepsilon$ for which the following estimate holds. For every $\lambda \geq 1$, one can find $k_0 = k_0(\lambda)$ such that for all $k \geq k_0$,*

$$\sup_{u\in[0,1]^d} \mathscr{B}_k(u,\lambda;s) \leq C_1\Phi(\lambda)\Theta(\lambda;s,\varrho;1)\prod_{m=1}^{k} p_m \tag{4.11}$$

where p_k denotes the selection probability given in (2.13)*, and s, ϱ, Φ are as in* (4.2)*.*

4.3. **Proof of Proposition 3.4 assuming Proposition 4.2.**

Proof. With all the notation in place, the only ingredient in this proof is the almost sure estimate (2.19) on P_k in terms of R_k, where R_k is as in (2.16). In view of the relation (4.7) linking $\mathscr{A}_k$ and $\mathscr{B}_k$, this estimate along with the conclusion (4.11) of Proposition 4.2 implies that for $\mathbb{P}^*$-almost every $\omega\in\Omega$, there exists a constant $C_2>0$ such that

$$\mathscr{A}_k(u,\lambda;s) \leq (C_2^{-1}R_k\delta_k^d)^{-1}C_1\Theta(\lambda;s,\varrho;\kappa)\prod_{m=1}^{k} p_m = C\Theta(\lambda;s,\varrho;\kappa) \tag{4.12}$$

for all large enough k. The last step above uses the relation $R_k\delta_k^d = p_1p_2\cdots p_k$, which can be easily deduced from (2.16) and (2.13). Since the right hand side of (4.12) is uniform in k, combining (4.12) with (4.10) leads to the conclusion of Proposition 3.4. □

5. Estimation of $\mathscr{B}_k$: Proof of Proposition 4.2

In view of the reduction carried out in Section 4, our goal now is to prove Proposition 4.2. We proceed to do so in this section, by rewriting $\mathscr{B}_k$ as a telescoping sum of centred random variables. Set

$$\mathscr{C}_k = \mathscr{C}_k(u,\lambda;s) := \sum_{\mathbf{i}_k} X_{k-1}(\mathbf{i}_{k-1})\big[Y_k(\mathbf{i}_k)-p_k\big]w_k(\mathbf{i}_k), \text{ so that} \tag{5.1}$$

$$\mathscr{B}_k := \mathscr{C}_k + p_k\mathscr{B}_{k-1}. \tag{5.2}$$

Here $p_k = \mathbb{E}(Y_k(\mathbf{i}_k))$ is as in (2.13) of Section 2.4. Iterating the recursion relation (5.2) yields

$$\mathscr{B}_k = \mathscr{C}_k + \sum_{\ell=1}^{k-1}\Big[\prod_{m=\ell+1}^{k} p_m\Big]\mathscr{C}_\ell + p_1p_2\cdots p_k\int_{[0.1]^d}\langle\lambda(u-v)\rangle^{-s}. \tag{5.3}$$

Thus $\mathscr{B}_k$ is a sum of $k+1$ terms, all of which are random except the last summand. We now proceed to estimate each term in the sum. We first show that the deterministic last term in (5.3) can be dominated by a suitable choice of Θ.

Lemma 5.1. *There exists an absolute constant $C = C_d > 0$ such that*

$$\sup_{u\in[0,1]^d}\int_{[0,1]^d}\langle\lambda(u-v)\rangle^{-s} \leq C_d\Theta(\lambda;s,d;\log(1+\lambda)),$$

where Θ *is as in* (4.1).

Proof. For $u \in [0,1]^d$, we make a change of variable $v \mapsto x = u - v$ in the integral to obtain

$$\int_{[0,1]^d} \langle \lambda(u-v)\rangle^{-s}\, dv \le \int_{|x|\le 2\sqrt{d}} \langle \lambda x\rangle^{-s}\, dx \le C_d \int_{r=0}^{2\sqrt{d}} \langle \lambda r\rangle^{-s} r^{d-1} dr$$
$$\le C_d \lambda^{-d} \int_0^{\lambda} \langle r\rangle^{-s} r^{d-1}\, dr.$$

The integrand in the second integral above is rotationally symmetric, hence the second inequality follows from a spherical change of coordinates $x = r\omega$, $r > 0$, $\omega \in \mathbb{S}^{d-1}$. The last inequality results from a scaling transformation $r \mapsto r/(2\lambda\sqrt{d})$. The last integral in the display above is precisely of the form dealt with in Lemma 4.1. Invoking (4.3) with $\vartheta = d$, $\rho = s$ therefore yields the desired conclusion. □

5.1. **An estimate for $\mathscr{C}_k$.** We now turn our attention to estimating $\mathscr{C}_k$. The main result here is Proposition 5.2, which will be proved in the next section. Even though $\mathscr{C}_k$ is a random quantity, the estimate that we seek will be deterministic and given in terms of a function $\Psi_k = \Psi_k(\lambda; s, d, \varepsilon)$ that we now define:

$$\Psi_k := \delta_k^d \prod_{m=1}^{k-1} p_m \times \Theta_k \quad \text{where} \tag{5.4}$$

$$\Theta_k := \left\{ \begin{array}{ll} \Theta\big(\delta_{k-1}^{-1}; 2s, \varrho; N_{k-1}^C\big)\Theta\big(\lambda\delta_{k-1}; 2s, 2d; \Phi(\lambda)\big) & \text{if } \lambda\delta_{k-1} > \lambda\delta_k > 1, \\ \Theta\big(\delta_{k-1}^{-1}; 2s, \varrho; N_{k-1}^C\big)\Theta\big(\lambda\delta_{k-1}; 2s, d; 1\big) & \text{if } \lambda\delta_k \le 1 < \lambda\delta_{k-1}, \\ \Theta(\lambda; 2s, \varrho; \Phi(\lambda)) & \text{if } \lambda\delta_{k-1} \le 1. \end{array} \right\} \tag{5.5}$$

Here Θ is as in (4.1), and C is an unspecified constant whose exact value may change from one occurrence to the next but which depends only on d, s and ε. As we will see, the value of C has no effect on the exponent α_p given in (3.13) and (3.15); it only affects β_p. The function Φ is of the form

$$\Phi(t) := \exp(C\sqrt{\log t}), \quad t \ge 1. \tag{5.6}$$

With this notation in place, our main estimate for $\mathscr{C}_k$ is the following.

Proposition 5.2. *Fix any two positive constants s and ρ. For $\mathbb{P}^*$-almost every $\omega \in \Omega$, there exists a constant $C > 0$ such that for all $\lambda \ge 1$ and $k \ge 1$,*

$$\sup_{u\in[0,1]^d} |\mathscr{C}_k(u; \lambda, s)| \le C \log(\lambda/\delta_k)^{\frac{1}{2}} \sqrt{\Psi_k}. \tag{5.7}$$

5.2. **Proof of Proposition 4.2, assuming Proposition 5.2.**

Proof. Let s and ϱ be as in (4.2). Dividing both sides of (5.3) by $p_1 p_2 \cdots p_k$ yields

$$\mathscr{B}_k \Big[\prod_{m=1}^{k} p_m\Big]^{-1} = \sum_{\ell=1}^{k} \Big[\prod_{m=1}^{\ell} p_m\Big]^{-1} \mathscr{C}_\ell + \int_{[0,1]^d} \langle \lambda(u-v)\rangle^{-s}\, dv$$
$$\le \sum_{\ell=1}^{k} \Big[\prod_{m=1}^{\ell} p_m\Big]^{-1} |\mathscr{C}_\ell| + \Theta(\lambda; s, d; \log\lambda), \tag{5.8}$$

$$\leq \sum_{\ell=1}^{k}\Big[\prod_{m=1}^{\ell} p_m\Big]^{-1} \log(\lambda/\delta_\ell)^{\frac{1}{2}}\sqrt{\Psi_\ell} + \Theta(\lambda; s, d; \log\lambda), \tag{5.9}$$

where the summands in the first term of (5.8) have been estimated using (5.7) of Proposition 5.2. The bound on the last term uses Lemma 5.1. We will prove shortly in Lemma 5.3 below, that for an absolute constant $C > 0$,

$$\sum_{\ell=1}^{k}\Big[\prod_{m=1}^{\ell} p_m\Big]^{-1} \log(\lambda/\delta_\ell)^{\frac{1}{2}}\sqrt{\Psi_\ell} \leq C\Theta(\lambda; s, \varrho; \Phi(\lambda)), \tag{5.10}$$

with Φ as in (5.6). Assuming this for the moment, and inserting (5.10) into (5.9), we obtain

$$\begin{aligned}\mathscr{B}_k\Big[\prod_{m=1}^{k} p_m\Big]^{-1} &\leq C\Big[\Theta(\lambda; s, \varrho; \Phi(\lambda)) + \Theta(\lambda; s, d; \log\lambda)\Big]\\ &\leq C\Theta(\lambda; s, \varrho; \Phi(\lambda)).\end{aligned}$$

The last inequality can be verified directly from the definition (4.1) of Θ, by comparing its values in the different regimes of s, ε and d. This completes the proof. □

5.3. **Estimating the sum in** (5.10)**.**

Lemma 5.3. *With s, ϱ, κ as in* (4.2)*, the estimate* (5.10) *holds.*

Proof. In view of the three-part description of Ψ_k given in (5.4) and (5.5), the sum on the left hand side of (5.10) can be decomposed into three sub-sums:

$$\sum_{\ell=1}^{k}\Big[\prod_{m=1}^{\ell} p_m\Big]^{-1} \log(\lambda/\delta_\ell)^{\frac{1}{2}}\sqrt{\Psi_\ell} = \mathscr{S}_1 + \mathscr{S}_2 + \mathscr{S}_3. \tag{5.11}$$

Here $\mathscr{S}_m$ denotes the sum over indices $\ell \in \mathscr{L}_m$ for $m = 1, 2, 3$, where

$$\begin{aligned}\mathscr{L}_1 &:= \{1 \leq \ell \leq k : \lambda\delta_\ell > 1\},\\ \mathscr{L}_2 &:= \{1 \leq \ell \leq k : \lambda\delta_\ell \leq 1 < \lambda\delta_{\ell-1}\},\\ \mathscr{L}_3 &:= \{1 \leq \ell \leq k : \lambda\delta_{\ell-1} \leq 1\}.\end{aligned}$$

In order to establish the claimed bound, it suffices to show that each $\mathscr{S}_m$ is bounded above by $\Phi(\lambda)\Theta(\lambda; s, \varrho; 1)$. The estimation accordingly splits into three steps.

Step 1: Estimating $\mathscr{S}_1$. The descriptions (5.4) and (5.5) of Ψ_k dictate that for indices $\ell \in \mathscr{L}_1$,

$$\begin{aligned}\Psi_\ell &= \delta_\ell^d\Big[\prod_{m=1}^{\ell-1} p_m\Big] \times \Theta(\delta_{\ell-1}^{-1}; 2s, \varrho; N_{\ell-1}^C)\Theta(\lambda\delta_{\ell-1}; 2s, 2d; \Phi(\lambda))\\ &= \delta_\ell^d \prod_{m=1}^{\ell-1} p_m \times \left\{\begin{array}{ll}\lambda^{-2s} & \text{if } s < \varrho/2,\\ \lambda^{-2s}\delta_{\ell-1}^{\varrho-2s}N_{\ell-1}^C & \text{if } \varrho/2 \leq s < d,\\ \lambda^{-2d}\delta_{\ell-1}^{\varrho-2d}N_{\ell-1}^C\Phi(\lambda) & \text{if } s \geq d.\end{array}\right\}\end{aligned} \tag{5.12}$$

We will insert the expressions for Ψ_ℓ obtained in (5.12) into the summand in (5.11) for various ranges of s and ϱ. Thus for $s < \varrho/2$, $\mathscr{S}_1$ reduces to

$$\begin{aligned}\mathscr{S}_1 &= \sum_{\ell\in\mathscr{L}_1} \delta_\ell^{\frac{d}{2}} \Big[\prod_{m=1}^{\ell-1} p_m\Big]^{-\frac{1}{2}} p_\ell^{-1}\lambda^{-s}\log(\lambda/\delta_\ell)^{\frac{1}{2}}\\ &\le C_{d,s,\varepsilon}\lambda^{-s}(\log\lambda)^{\frac{1}{2}} \le C_{d,s,\varepsilon}\Phi(\lambda)\Theta(\lambda;s,\varrho;1).\end{aligned}$$

For $\varrho/2 \le s < d$, inserting (5.12) into (5.11) yields

$$\begin{aligned}\mathscr{S}_1 &= \sum_{\ell\in\mathscr{L}_1} \delta_\ell^{\frac{d}{2}} \Big[\prod_{m=1}^{\ell-1} p_m\Big]^{-\frac{1}{2}} p_\ell^{-1}\log(\lambda/\delta_\ell)^{\frac{1}{2}}\lambda^{-s}\delta_{\ell-1}^{\frac{\varrho}{2}-s}N_{\ell-1}^C\\ &\le C_{d,s}\lambda^{-s}\sum^{[1]}\delta_\ell^{\varrho-s}\log(\lambda/\delta_\ell)^{\frac{1}{2}}N_\ell^C\\ &\le C_{d,s}\lambda^{-s}\times\left\{\begin{array}{ll}(\log\lambda)^{\frac{1}{2}} & \text{if } \frac{\varrho}{2}\le s<\varrho,\\ \lambda^{s-\varrho}\Phi(\lambda) & \text{if } \varrho\le s<d\end{array}\right\}\\ &\le C_{d,s,\varepsilon}\Phi(\lambda)\Theta(\lambda;s,\varrho;1).\end{aligned}$$

For $s \ge d$, the same procedure leads to

$$\begin{aligned}\mathscr{S}_1 &= \lambda^{-d}\sum_{\ell\in\mathscr{L}_1}\delta_\ell^{\frac{d}{2}}\big[\prod_{m=1}^{\ell-1}p_m\big]^{-\frac{1}{2}}p_\ell^{-1}\log(\lambda/\delta_\ell)^{\frac{1}{2}}\delta_{\ell-1}^{\frac{\varrho}{2}-d}N_{\ell-1}^C\Phi(\lambda)\\ &\le C_{d,s,\varepsilon}\lambda^{-d}\sum_{\ell\in\mathscr{L}_1}\delta_\ell^{-d\varepsilon}N_\ell^C\log(\lambda/\delta_\ell)^{\frac{1}{2}}\\ &\le C_{d,s,\varepsilon}\lambda^{-\varrho}\Phi(\lambda)\le C_{d,s,\varepsilon}\Phi(\lambda)\Theta(\lambda;s,\varrho;1).\end{aligned}$$

This completes the proof for $\mathscr{S}_1$.

Step 2: Estimating $\mathscr{S}_2$. The number of indices ℓ obeying $\lambda\delta_\ell \le 1 < \lambda\delta_{\ell-1}$ is exactly one, so the sub-sum $\mathscr{S}_2$ is in fact a single term, namely

$$\begin{aligned}\mathscr{S}_2 &= \Big[\prod_{m=1}^{\ell}p_m\Big]^{-1}\log(\lambda/\delta_\ell)^{\frac{1}{2}}\sqrt{\Psi_\ell}\\ &= \delta_\ell^{\frac{d}{2}}\Big[\prod_{m=1}^{\ell-1}p_m\Big]^{-\frac{1}{2}}p_\ell^{-1}\log(\lambda/\delta_\ell)^{\frac{1}{2}}\Big[\Theta(\delta_{\ell-1}^{-1};2s,\varrho;N_{\ell-1}^C)\Theta(\lambda\delta_{\ell-1};2s,d;1)\Big]^{\frac{1}{2}}.\end{aligned}$$

Let us pause for a moment to observe that in this case

$$\delta_\ell = O(\lambda^{-1}),\quad \delta_{\ell-1} = O(\lambda^{-1}\Phi(\lambda)),\quad \text{and hence}\quad \lambda\delta_{\ell-1} = O(N_{\ell-1}^C).$$

We will use these facts without further reference in the remainder of this scenario. As in the previous case, we evaluate this term for different values of s and ϱ. For $s < \varrho/2$,

$$\mathscr{S}_2 = \delta_\ell^{\frac{d}{2}}\Big[\prod_{m=1}^{\ell-1}p_m\Big]^{-\frac{1}{2}}p_\ell^{-1}\log(\lambda/\delta_\ell)^{\frac{1}{2}}\big[\delta_{\ell-1}^{2s}(\lambda\delta_{\ell-1})^{-2s}\big]^{\frac{1}{2}}$$

$$= \lambda^{-s}\delta_\ell^{\frac{\varrho}{2}} N_{\ell-1}^C \log(\lambda/\delta_\ell)^{\frac12} \le C_{d,s,\varepsilon}\lambda^{-s}\Phi(\lambda). \tag{5.13}$$

For $\varrho/2 \le s < d/2$ (which occurs if and only if $\varepsilon > 0$),

$$\begin{aligned}\mathscr{S}_2 &= \delta_\ell^{\frac{d}{2}}\Big[\prod_{m=1}^{\ell-1} p_m\Big]^{-\frac12} p_\ell^{-1}\log(\lambda/\delta_\ell)^{\frac12}\big[\delta_{\ell-1}^{\varrho} N_{\ell-1}^C(\lambda\delta_{\ell-1})^{-2s}\big]^{\frac12}\\ &\le \lambda^{-s}\delta_\ell^{\varrho-s}N_{\ell-1}^C \le C_{d,s,\varepsilon}\lambda^{-\varrho}\Phi(\lambda). \end{aligned}\tag{5.14}$$

For $s \ge d/2$,

$$\begin{aligned}\mathscr{S}_2 &= \delta_\ell^{\frac{d}{2}}\Big[\prod_{m=1}^{\ell-1} p_m\Big]^{-\frac12} p_\ell^{-1}\log(\lambda/\delta_\ell)^{\frac12}\big[\delta_{\ell-1}^{\varrho}(\lambda\delta_{\ell-1})^{-d} N_{\ell-1}^C\big]^{\frac12}\\ &\le \lambda^{-\frac{d}{2}}\delta_\ell^{\varrho-\frac{d}{2}}\log(\lambda/\delta_\ell)^{\frac12}N_{\ell-1}^C \le C_{d,s,\varepsilon}\lambda^{-\varrho}\Phi(\lambda). \end{aligned}\tag{5.15}$$

Combining the final estimates in (5.13), (5.14) and (5.15), we find that

$$\mathscr{S}_2 \le C_{d,s,\varepsilon}\times\Phi(\lambda)\times\begin{Bmatrix}\lambda^{-s} & \text{if } s<\varrho/2\\ \lambda^{-\varrho} & \text{if } s\ge\varrho/2\end{Bmatrix} \le C_{d,s,\varepsilon}\Phi(\lambda)\Theta(\lambda;s,\varrho;1).$$

The verification of the last inequality is left to the reader.

Step 3: Estimating $\mathscr{S}_3$. In this case,

$$\sqrt{\Psi_\ell} = \delta_\ell^{\frac{d}{2}}\Big[\prod_{m=1}^{\ell-1} p_m\Big]^{\frac12}\sqrt{\Theta(\lambda;2s,\varrho;\Phi(\lambda))}.$$

Inserting this into the expression for $\mathscr{S}_3$ yields

$$\begin{aligned}\mathscr{S}_3 &= \sum_{\ell\in\mathscr{L}_3}\delta_\ell^{\frac{d}{2}}\Big[\prod_{m=1}^{\ell-1} p_m\Big]^{-\frac12}\log(\lambda/\delta_\ell)^{\frac12}\times\begin{Bmatrix}\lambda^{-s} & \text{if } s<\varrho/2,\\ \lambda^{-\varrho/2} & \text{if } s\ge\varrho/2\end{Bmatrix}\\ &\le \begin{Bmatrix}\lambda^{-s}\sum^{[3]}\delta_\ell^{\frac{\varrho}{2}}\log(\lambda/\delta_\ell)^{\frac12}N_{\ell-1}^C & \text{if } s<\varrho/2,\\ \lambda^{-\frac{\varrho}{2}}\Phi(\lambda)\sum^{[3]}\delta_\ell^{\frac{\varrho}{2}}\log(\lambda/\delta_\ell)^{\frac12}N_{\ell-1}^C & \text{if } s\ge\varrho/2\end{Bmatrix}\\ &\le C_{d,s,\varepsilon}\times\Phi(\lambda)\times\begin{Bmatrix}\lambda^{-s} & \text{if } s<\varrho/2,\\ \lambda^{-\varrho} & \text{if } s\ge\varrho/2,\end{Bmatrix}\\ &\le C_{d,s,\varepsilon}\Phi(\lambda)\Theta(\lambda;s,\varrho;1),\end{aligned}$$

where the last step follows from the definition (4.1) of Θ. This completes the proof of the lemma. □

6. Estimation of $\mathscr{C}_k$: Proof of Proposition 5.2

Summarizing the situation thus far, we have reduced the proof of Theorem 1.3 to that of Proposition 5.2. We complete the latter proof in this section, modulo two propositions based on large deviation inequalities that will be proved subsequently.

We need the following auxiliary quantity:

$$\mathscr{D}_k = \mathscr{D}_k(u,\lambda;s) := \sum_{\mathbf{i}_k} X_{k-1}(\mathbf{i}_{k-1})\big(w_k(\mathbf{i}_k)\big)^2 = \sum_{\mathbf{i}_{k-1}} X_{k-1}(\mathbf{i}_{k-1}) \sum_{\bar{i}_k} (w_k(\mathbf{i}_k))^2, \tag{6.1}$$

which will be used to bound $\mathscr{C}_k$. We will also need a fine discretization of the frequency scale λ and the spatial parameter u. With this in mind, let us decompose the range of the frequency parameter λ into countably many pieces using

$$\mathbf{\Lambda}_j := [\delta_j^{-1}, \delta_{j+1}^{-1}], \quad \text{so that} \quad [\delta_1^{-1}, \infty) = \bigcup_{j=1}^{\infty} \mathbf{\Lambda}_j. \tag{6.2}$$

For a constant M soon to be specified, we fix maximal $\frac{1}{10}(\delta_k\delta_{j+1})^M$-separated sets

$$\mathbb{U}_{jk} = \mathbb{U}_{jk}[M] \subseteq [0,1]^d \quad \text{and} \quad \mathbf{\Lambda}_{jk} = \mathbf{\Lambda}_{jk}[M] \subseteq \mathbf{\Lambda}_j, \tag{6.3}$$

so that

$$\#(\mathbb{U}_{jk}) \le 10^d \delta_k^{-Md}\delta_{j+1}^{-Md}, \quad \text{and} \quad \#(\mathbf{\Lambda}_{jk}) \le 10\delta_{j+1}^{-1-M}\delta_k^{-M}. \tag{6.4}$$

The relevance of $\mathscr{D}_k$, $\mathbb{U}_{jk}$ and $\mathbf{\Lambda}_{jk}$ in the estimation of $\mathscr{C}_k$ is clarified in the following two propositions.

Proposition 6.1. *Any choice of a large constant $R \ge 1$ permits the choice of an absolute constant $M_R = M(R,d)$ with the following property.*

For $\mathbb{P}^*$-almost every $\omega \in \Omega$, there exists a constant $C = C(R,\omega) > 0$ such that for each $j,k \ge 1$, $\lambda \in \mathbf{\Lambda}_j$ and $u \in [0,1]^d$, one can find $u' \in \mathbb{U}_{jk}[M_R]$ and $\lambda' \in \mathbf{\Lambda}_{jk}[M_R]$ obeying the relation

$$|\mathscr{C}_k(u,\lambda;s)| \le C|\log(\delta_{j+1}\delta_k)|^{\frac{1}{2}}\sqrt{\mathscr{D}_k(u',\lambda';s)} + (\delta_{j+1}\delta_k)^R. \tag{6.5}$$

Here $\mathscr{C}_k, \mathscr{D}_k$ are as in (5.1) *and* (6.1) *respectively.*

Proposition 6.2. *For $\mathbb{P}^*$-almost every $\omega \in \Omega$, there exists a constant $C = C(\omega,d,s,\varepsilon) > 0$ such that for all indices $j,k \ge 1$, all $\lambda \in \mathbb{L}_j$ and $u \in [0,1]^d$, the quantity $\mathscr{D}_k$ defined in* (6.1) *admits the following bound:*

$$\mathscr{D}_k(u,\lambda;s) \le C\Psi_k(\lambda;s,d,\varepsilon), \tag{6.6}$$

with Ψ_k as in (5.4).

6.1. **Proof of Proposition 5.2, assuming Propositions 6.1 and 6.2.**

Proof. Let R be an absolute constant depending only on d and s such that

$$(\delta_{j+1}\delta_k)^R \le \inf\Big\{\sqrt{\Psi_k(\lambda;s,d,\varepsilon)} : \lambda \in \mathbf{\Lambda}_j\Big\}.$$

In fact any $R > 100(d+s)$ will suffice. The desired estimate (5.7) is then obtained by simply combining the two inequalities (6.5) with (6.6). □

7. Estimation of $\mathscr{C}_k$ via a large deviation inequality

7.1. Proof of Proposition 6.1.

Proof. The argument relies on two key steps. The first is a suitably fine discretization of the parameters u and λ, as a result of which the estimation of the supremum of $\mathscr{C}_k(u,\lambda;s)$ is reduced to its evaluation at finitely many points. This step has been carried out in Proposition 7.2 below. The second is an application of Azuma's inequality, quoted in Theorem 12.4, on each of the finitely many $\mathscr{C}_k(u,\lambda;s)$ thus obtained. The details of this are in Lemma 7.1.

Assuming Lemma 7.1 and Proposition 7.2 for the moment, the proof is completed as follows. Proposition 7.2 dictates that for every $R \geq 1$, there exists a constant $M_R > 0$ such that for every $(u,\lambda),(u',\lambda') \in [0,1]^d \times \mathbf{\Lambda}_j$,

$$|\mathscr{C}_k(u,\lambda;s) - \mathscr{C}_k(u',\lambda';s)| \leq (\delta_{j+1}\delta_k)^R \quad \text{provided} \quad |(u,\lambda)-(u',\lambda')| \leq (\delta_k\delta_{j+1})^{M_R}. \tag{7.1}$$

In view of the separation condition on $\mathbb{U}_{jk}$ and $\mathbf{\Lambda}_{jk}$ described earlier in this section in the lead up to (6.3), this means that for every $(u,\lambda) \in [0,1]^d \times \mathbf{\Lambda}_j$, there exists $(u',\lambda') \in \mathbb{U}_{jk}[M_R] \times \mathbf{\Lambda}_{jk}[M_R]$ such that

$$|\mathscr{C}_k(u,\lambda;s) - \mathscr{C}_k(u',\lambda';s)| \leq (\delta_{j+1}\delta_k)^R.$$

Combining this with the conclusion (7.2) of Lemma 7.1, we arrive at the estimate

$$\begin{aligned} |\mathscr{C}_k(u,\lambda;s)| &\leq |\mathscr{C}_k(u',\lambda';s)| + (\delta_{j+1}\delta_k)^R \\ &\leq C|\log(\delta_{j+1}\delta_k)|^{\frac12}\sqrt{\mathscr{D}_k(u',\lambda';s)} + (\delta_{j+1}\delta_k)^R, \end{aligned}$$

which is the claimed inequality (6.5). □

Lemma 7.1. *Fix $M \geq 1$. For $\mathbb{P}^*$-almost every ω, there exists a constant $C = C(M,\omega) > 0$ such that for all $j,k \geq 1$ and $(u,\lambda) \in \mathbb{U}_{jk}[M] \times \mathbf{\Lambda}_{jk}[M]$,*

$$|\mathscr{C}_k(u,\lambda;s)| \leq C|\log(\delta_{j+1}\delta_k)|^{\frac12}\sqrt{\mathscr{D}_k(u,\lambda;s)}. \tag{7.2}$$

Proof. We follow a reasoning similar to the proof of (2.18) in Lemma 2.3. For a large absolute constant B depending only on M and soon to be specified, we define an event $\mathtt{S}_{jk}$ as follows,

$$\mathtt{S}_{jk}^c := \left\{ \omega \in \Omega \,\middle|\, \begin{array}{l} \text{there exists } (u,\lambda) \in \mathbb{U}_{jk}[M] \times \mathbf{\Lambda}_{jk}[M] \text{ such that} \\ |\mathscr{C}_k(u,\lambda;s)| > B|\log(\delta_{j+1}\delta_k)|^{\frac12}\sqrt{\mathscr{D}_k(u,\lambda;s)} \end{array} \right\}.$$

This event may be rewritten as

$$\mathtt{S}_{jk}^c = \bigcup_{u,\lambda} \big[\mathtt{S}_{jk}(u,\lambda;s)\big]^c \quad \text{where} \tag{7.3}$$

$$\big[\mathtt{S}_{jk}(u,\lambda;s)\big]^c := \left\{ \omega \in \Omega : |\mathscr{C}_k(u,\lambda;s)| > B|\log(\delta_{j+1}\delta_k)|^{\frac12}\sqrt{\mathscr{D}_k(u,\lambda;s)} \right\}, \tag{7.4}$$

and the union in (7.3) takes place over all tuples $(u,\lambda) \in \mathbb{U}_{jk}[M] \times \mathbf{\Lambda}_{jk}[M]$. It is possible for the set in (7.4) to be empty for certain choices of (u,λ). We will shortly show that

$$\sum_{j,k=1}^{\infty} \mathbb{P}^*(\mathtt{S}_{jk}^c) < \infty. \tag{7.5}$$

Once this is proved, the Borel-Cantelli lemma would imply that $\mathbb{P}^*$-almost surely, the event $\mathtt{S}_{jk}$ holds for all but finitely many j and k. More precisely, for $\mathbb{P}^*$-almost every $\omega \in \Omega$, there exist $j_0(\omega)$ and $k_0(\omega)$ such that except for $(j,k) \in [1, j_0(\omega)] \times [1, k_0(\omega)]$,

$$|\mathscr{C}_k(u,\lambda;s)| \le B|\log(\delta_{j+1}\delta_k)|^{\frac{1}{2}}\sqrt{\mathscr{D}_k(u,\lambda;s)} \text{ for all } (u,\lambda) \in \mathbb{U}_{jk}[M] \times \mathbf{\Lambda}_{jk}[M].$$

After possibly adjusting the value of C in order to accommodate the remaining finitely many values of j and k, we can ensure that (7.2) holds for all j and k, with

$$C = B + \max\left\{\frac{|\mathscr{C}_k(u,\lambda;s)|}{|\log(\delta_{j+1}\delta_k)|^{\frac{1}{2}}\sqrt{\mathscr{D}_k(u,\lambda;s)}} : (u,\lambda) \in \mathbb{U}_{jk} \times \mathbf{\Lambda}_{jk},\ 1 \le j \le j_0, 1 \le k \le k_0\right\}.$$

We pause for a moment to observe that the limiting set E is nonempty on the support of $\mathbb{P}^*$; hence by (2.7), $P_k > 0$ for every $k \ge 1$. Since $w_k(\mathbf{i}_k)$ is always strictly positive, this implies that the quantity $\mathscr{D}_k(u,\lambda;s)$ defined in (6.1) is also strictly positive for every (u,λ). As a result, the constant C defined above is also positive and finite.

It remains to prove (7.5). Let $\mathcal{F}_k$ denote the σ-algebra generated by $\mathbf{Y}_1, \cdots, \mathbf{Y}_k$. Let us write

$$\mathbb{P}^*(\mathtt{S}_{jk}^c) = \frac{\mathbb{P}(\mathtt{T}_{jk})}{\mathbb{P}(E \ne \emptyset)} = \frac{\mathbb{E}(\mathbb{P}(\mathtt{T}_{jk}|\mathcal{F}_{k-1}))}{\mathbb{P}(E \ne \emptyset)} \quad \text{with } \mathtt{T}_{jk} = \mathtt{S}_{jk}^c \cap \{E \ne \emptyset\}.$$

We observe that the denominator is a nonzero constant independent of j,k by Lemma 2.2. Thus for (7.5), it suffices to show that $\mathbb{P}(\mathtt{T}_{jk}|\mathcal{F}_{k-1})$ is bounded above by a deterministic constant that is summable in j,k. We estimate $\mathbb{P}(\mathtt{T}_{jk}|\mathcal{F}_{k-1})$ using Azuma's inequality, quoted in Theorem 12.4 below. Conditioning on $\mathcal{F}_{k-1}$, we observe that for each (u,λ), the quantity $\mathscr{C}_k$ as defined by (5.1) is a sum of independent, centred random variables, and in particular a martingale. More precisely, in the notation of Theorem 12.4, one has

$$U_0 = 0, \quad m = M_k, \quad U_m = U_m - U_0 = \mathscr{C}_k(u,\lambda;s), \quad c_{\mathbf{i}_k} = X_{k-1}(\mathbf{i}_{k-1})w_k(\mathbf{i}_k).$$

Setting

$$t = B|\log(\delta_{j+1}\delta_k)|^{\frac{1}{2}}\Big(\sum_{\mathbf{i}_k} c_{\mathbf{i}_k}^2\Big)^{\frac{1}{2}}$$

in (12.17), we obtain for each choice of (u,λ),

$$\mathbb{P}(\mathtt{T}_{jk}(u,\lambda;s)|\mathcal{F}_{k-1}) \le 2\exp\Big(-\frac{t^2}{2\sum_{\mathbf{i}_k} c_{\mathbf{i}_k}^2}\Big) \le 2(\delta_{j+1}\delta_k)^{\frac{B^2}{2}}, \qquad \text{where}$$
$$\mathtt{T}_{jk}(u,\lambda;s) := [\mathtt{S}_{jk}(u,\lambda;s)]^c \cap \{E \ne \emptyset\},$$

with $\mathtt{S}_{jk}(u,\lambda;s)$ as in (7.4). Since $\mathtt{T}_{jk}$ is the union of the events $\mathtt{T}_{jk}(u,\lambda;s)$ for $(u,\lambda) \in \mathbb{U}_{jk}[M] \times \mathbf{\Lambda}_{jk}[M]$, the trivial bound gives

$$\begin{aligned}\mathbb{P}(\mathtt{T}_{jk}|\mathcal{F}_{k-1}) &\le 2(\delta_{j+1}\delta_k)^{\frac{B^2}{2}} \times \#(\mathbb{U}_{jk}[M]) \times \#(\mathbf{\Lambda}_{jk}[M]) \\ &\le C_d(\delta_{j+1}\delta_k)^{\frac{B^2}{2}-1-M(d+1)},\end{aligned}$$

where the last step follows from (6.4). For an absolute large constant B obeying $B^2 > 2 + 2M(d+1)$, the right hand side above is a deterministic constant summable in j and k, completing the proof. □

7.2. **Discretization.**

Proposition 7.2. *Let the parameters d, s be as in Section 4.1. Then for any large constant $R \geq 1$, there exists $M = M(R, d, n, s)$ with the following property. For every $j, k \geq 1$, $u, u' \in [0,1]^d$, $\lambda, \lambda' \in \mathbf{\Lambda}_j = [\delta_j^{-1}, \delta_{j+1}^{-1}]$ with*

$$|u - u'| \leq (\delta_{j+1}\delta_k)^M, \qquad |\lambda - \lambda'| \leq (\delta_{j+1}\delta_k)^M \tag{7.6}$$

and every $\omega \in \Omega$, the function $\mathscr{C}_k = \mathscr{C}_{k,\omega}$ given by (5.1) obeys the estimate

$$\Big|\mathscr{C}_k(u;\lambda,s) - \mathscr{C}_k(u';\lambda',s)\Big| \leq (\delta_{j+1}\delta_k)^R. \tag{7.7}$$

Proof. The mean value theorem shows that the function $z \mapsto \langle z\rangle^{-s}$ is uniformly Lipschitz on $\mathbb{R}^d$, i.e., there exists an absolute constant $B = B_{s,d}$ such that

$$\big|\langle z\rangle^{-s} - \langle z'\rangle^{-s}\big| \leq B|z - z'| \quad \text{for all } z, z' \in \mathbb{R}^d. \tag{7.8}$$

This means that for w_k defined as in (4.9) and for any multi-index $\mathbf{i}_k$,

$$\begin{aligned}
&\big|w_k(\mathbf{i}_k;u,\lambda,s) - w_k(\mathbf{i}_k;u',\lambda',s)\big| \\
&\qquad \leq \big|w_k(\mathbf{i}_k;u,\lambda,s) - w_k(\mathbf{i}_k;u',\lambda,s)\big| + \big|w_k(\mathbf{i}_k;u',\lambda,s) - w_k(\mathbf{i}_k;u',\lambda',s)\big| \\
&\qquad \leq \int_{Q(\mathbf{i}_k)} \big|\big[\langle\lambda(u-v)\rangle^{-s} - \langle\lambda(u'-v)\rangle^{-s}\big]\,dv\big| \\
&\qquad\qquad\qquad + \int_{Q(\mathbf{i}_k)} \big|\big[\langle\lambda(u'-v)\rangle^{-s} - \langle\lambda'(u'-v)\rangle^{-s}\big]\,dv\big| \\
&\qquad \leq \int_{Q(\mathbf{i}_k)} B|\lambda(u-v) - \lambda(u'-v)|\,dv + \int_{Q(\mathbf{i}_k)} B|\lambda(u-v) - \lambda'(u'-v)|\,dv \\
&\qquad \leq B\delta_k^d(\lambda|u-u'| + |\lambda-\lambda'||u'|) \\
&\qquad \leq B\delta_k^d(\delta_{j+1}^{-1} + \sqrt{d})(\delta_{j+1}\delta_k)^M \leq 2B\delta_k^d(\delta_{j+1}\delta_k)^{M-1}.
\end{aligned} \tag{7.9}$$

The third inequality above is obtained by using the estimate (7.8) to estimate the integrands, once with $(z, z') = (\lambda(u-v), \lambda(u'-v))$, and once with $(z, z') = (\lambda(u'-v), \lambda'(u'-v))$. Substituting the estimate in (7.9) into the expression (5.1) for $\mathscr{C}_k$ yields the estimate:

$$\begin{aligned}
\big|\mathscr{C}_k(u,\lambda;s) - \mathscr{C}_k(u',\lambda';s)\big| &\leq \sum_{\mathbf{i}_k} X_{k-1}(\mathbf{i}_{k-1})|Y_k(\mathbf{i}_k) - p_k|\big|w_k(\mathbf{i}_k;u,\lambda,s) - w_k(\mathbf{i}_k;u',\lambda',s)\big| \\
&\leq 2B\delta_k^d(\delta_{j+1}\delta_k)^{M-1}\sum_{\mathbf{i}_k} X_{k-1}(\mathbf{i}_{k-1}) \\
&\leq 2B(\delta_{j+1}\delta_k)^{M-1},
\end{aligned}$$

since $\sum_{\mathbf{i}_k} X_{k-1}(\mathbf{i}_{k-1}) = P_{k-1}N_k^d \leq \delta_k^{-d}$. We now choose $M \gg R$ so that $\delta_1^{M-R-1} \leq (2B)^{-1}$. This ensures that the rightmost quantity in the displayed sequence of inequalities above is bounded by $(\delta_{j+1}\delta_k)^R$, as claimed. □

8. Estimation of deterministic weight functions

It remains to prove Proposition 6.2. This is our main objective for this section or the next. In preparation, let us recall from (6.1) the definition of the random quantity $\mathscr{D}_k$. The right hand side of (6.1) is a linear combination of the binary random variables $X_{k-1}(\mathbf{i}_{k-1})$ weighted by a deterministic weight function depending on k. Each deterministic weight is a short sum of the form

$$\sum_{\bar{i}_k} \big(w_k(\mathbf{i}_k)\big)^2 \tag{8.1}$$

involving $w_k = w_k(\mathbf{i}_k; u, \lambda, s)$, which in turn has been defined in (4.9). Our first order of business is to obtain good quantitative bounds on this weight function. The lemmas in this section provide these bounds, which will be used in the next section towards the proof of Proposition 6.2.

Lemma 8.1. *For every $u \in [0,1]^d$, $\lambda \geq 1$ and $k \geq 1$ and every multi-index $\mathbf{i}_{k-1} \in \mathbb{I}(k-1, d)$, the following estimate holds:*

$$\sum_{\bar{i}_k} (w_k(\mathbf{i}_{k-1}, \bar{i}_k))^2 \leq \delta_k^d \delta_{k-1}^d \langle \lambda \Delta(u, \mathbf{i}_{k-1}) \rangle^{-2s}, \tag{8.2}$$

where the summation index $\bar{i}_k$ ranges over $\mathbb{Z}^d_{N_k}$ and $\Delta(u, \mathbf{i}_k)$ denotes the distance of the point u from the cube $Q(\mathbf{i}_k)$.

Proof. Let us fix u, λ, k and $\mathbf{i}_{k-1}$. By the nesting property of the cubes $Q(\mathbf{i}_k)$, we know that $Q(\mathbf{i}_k) \subseteq Q(\mathbf{i}_{k-1})$ for every multi-index $\mathbf{i}_k$ of the form $\mathbf{i}_k = (\mathbf{i}_{k-1}, \bar{i}_k)$. Thus, by definition of Δ, any $v \in Q(\mathbf{i}_{k-1})$ satisfies the bound $|u - v| \geq \Delta(u, \mathbf{i}_{k-1})$. We substitute this bound into the expression (4.9) of w_k and recall that each cube $Q(\mathbf{i}_k)$ has sidelength δ_k and hence Lebesgue volume δ_k^d. This leads to the following pointwise bound on w_k:

$$w_k(\mathbf{i}_k) = w_k(\mathbf{i}_k; u, \lambda, s) = \int_{Q(\mathbf{i}_k)} \langle \lambda(u - v) \rangle^{-s} \, dv \leq \delta_k^d \langle \lambda \Delta(u, \mathbf{i}_{k-1}) \rangle^{-s}.$$

Since the right hand side above is independent of $\bar{i}_k \in \mathbb{Z}^d_{N_k}$, squaring and summing in $\bar{i}_k$ yields

$$\begin{aligned} \sum_{\bar{i}_k} (w_k(\mathbf{i}_k))^2 &\leq \delta_k^{2d} \langle \lambda \Delta(u, \mathbf{i}_{k-1}) \rangle^{-2s} N_k^d \\ &\leq \delta_k^d \delta_{k-1}^d \langle \lambda \Delta(u, \mathbf{i}_{k-1}) \rangle^{-2s}, \end{aligned}$$

where the last inequality uses the fact that $\delta_k N_k = \delta_{k-1}$, as seen from (2.1). This is the desired inequality (8.2). □

If u is close to $Q(\mathbf{i}_{k-1})$, and for certain regimes of λ, k, s and ϱ, the estimate in Lemma 8.1 can be improved. The next lemma quantifies this improvement. Set

$$\zeta_k := \max(\lambda \delta_k, 1), \tag{8.3}$$

and define a function $\Xi_k = \Xi_k(\lambda, \zeta_k, \zeta_{k-1}; s, d)$ as follows:

$$\Xi_k := \delta_k^d \delta_{k-1}^d \times \left\{ \begin{array}{l} 1 \text{ if } \zeta_{k-1} = 1, \text{ i.e., } \lambda\delta_{k-1} \le 1 \text{ and for any } s \text{ and } d, \\ \Theta(\zeta_{k-1}; 2s, d; \log(1+\zeta_{k-1})) \text{ if } \lambda\delta_k \le 1 < \lambda\delta_{k-1}, \\ \zeta_{k-1}^{-2s} \times \left\{ \begin{array}{ll} 1 & \text{if } s < \frac{d}{2} \text{ and } \zeta_k = \lambda\delta_k > 1, \\ \log N_k & \text{if } s = \frac{d}{2} \text{ and } \zeta_k = \lambda\delta_k > 1, \\ N_k^{-d+2s} & \text{if } \frac{d}{2} < s < d \text{ and } \zeta_k = \lambda\delta_k > 1, \end{array} \right\} \\ \zeta_k^{-d}\zeta_{k-1}^{-d} \log(1+\zeta_k)^2 \text{ if } d \le s \text{ and } \zeta_k = \lambda\delta_k > 1. \end{array} \right\} \tag{8.4}$$

Here the function Θ is as in (4.1). The reason for defining Ξ_k as above will emerge shortly, as we estimate each short sum of the form (8.1) over a variety of regimes in $d, s, \lambda, \delta_k, \delta_{k-1}$. Its immediate relevance is that it dominates this sum, as shown in the following lemma.

Lemma 8.2. *Given any constant $C_1 > 0$, there exists another constant $C_2 = C_2(C_1, d, s)$ with the following property.*

For any $u \in [0,1]^d$ and $k \ge 1$ such that $\Delta(u, \mathbf{i}_{k-1}) \le C_1\delta_{k-1}$, we have the estimate

$$\sum_{\bar{i}_k} (w_k(\mathbf{i}_{k-1}, \bar{i}_k))^2 \le C_2 \Xi_k, \tag{8.5}$$

where the function Ξ_k is as in (8.4).

The proof of the lemma above uses of a set of tools that we lay out in the following subsection.

8.1. **Refined estimation of the weights on cubes near u.** For any multi-index $\mathbf{i}_{k-1} \in \mathbb{I}(k-1, d)$, define a collection of cubes $\mathcal{Q}$ given by

$$\mathcal{Q} = \mathcal{Q}[\mathbf{i}_{k-1}] := \{Q(\mathbf{i}_{k-1}) + \mathbf{t} : \mathbf{t} = (t_1, \cdots, t_d) \in \mathbb{Z}^d,\ |t_i| \le C_1 \text{ for } 1 \le i \le d\}.$$

The number of cubes in $\mathcal{Q}$ is at most $(2C_1+1)^d$; further, $\mathcal{Q}$ has the property that any $u \in [0,1]^d$ obeying $\Delta(u, \mathbf{i}_{k-1}) \le C_1\delta_{k-1}$ must lie in at least one of the cubes in $\mathcal{Q}$. Fix one such point u, and let $\mathbf{i}_{k-1}^* \in \mathbb{I}(k-1, d)$ and $\mathbf{i}_k^* \in \mathbb{I}(k, d)$ denote multi-indices depending on u such that

$$u \in Q(\mathbf{i}_k^*) \subseteq Q(\mathbf{i}_{k-1}^*) \in \mathcal{Q}.$$

Let us consider the union Q^* of $Q(\mathbf{i}_k^*)$ along with its adjacent cubes. Thus Q^* is a cube of sidelength $3\delta_k$ containing $Q(\mathbf{i}_k^*)$ and axis-parallel to it, with $\text{dist}(u, \partial Q^*) \ge \delta_k$.

We cover $Q(\mathbf{i}_{k-1})$ using axis-parallel translates of Q^* of the form $Q^* + 3\mathbf{m}\delta_k$, $\mathbf{m} \in \mathbb{Z}^d$. Thus, we are able to ensure the following geometric properties:

(i) The possible values of the integer vector $\mathbf{m}$ needed for the covering are contained in $[-C_1N_k/3, C_1N_k/3]^d \cap \mathbb{Z}^d$.
(ii) Every cube $Q(\mathbf{i}_k) = Q(\mathbf{i}_{k-1}, \bar{i}_k)$ is contained in $Q^* + 3\mathbf{m}\delta_k$ for some $\mathbf{m}$ in the above range. The covering is essentially optimal in the following sense: for every fixed $\mathbf{m}$, we have $\#\{\bar{i}_k : Q(\mathbf{i}_k) \subseteq Q^* + 3\mathbf{m}\delta_k\} \le 3^d$.
(iii) If $v \in Q^* + 3\mathbf{m}\delta_k$ for some $\mathbf{m} \ne \mathbf{0}$, then $|u - v| \ge c_d|\mathbf{m}|\delta_k$.

8.2. **Proof of Lemma 8.2.**

Proof. Combining the observations (i)-(iii) in the previous subsection, we are led to the following estimate:

$$\sum_{\bar{i}_k}\big(w_k(\mathbf{i}_{k-1},\bar{i}_k)\big)^2 \le C_d \sum_{\mathbf{m}}{}' \Big[\int_{Q^*+3\mathbf{m}\delta_k} \langle\lambda(u-v)\rangle^{-s}\,dv\Big]^2 \quad \text{(by (i) and (ii) above)}$$

$$\le C_{d,s} \sum_{|\mathbf{m}|\le 10d}{}' \Big[\int_{Q^*+3\mathbf{m}\delta_k} \langle\lambda(u-v)\rangle^{-s}\,dv\Big]^2 + \sum_{|\mathbf{m}|>10d}{}' \langle\lambda\delta_k\mathbf{m}\rangle^{-2s}\delta_k^{2d} \quad \text{(using (iii) above)} \tag{8.6}$$

$$\le C_{d,s}(\mathscr{W}_1+\mathscr{W}_2). \tag{8.7}$$

The notation $\sum'$ appearing in the first and second steps above indicates that the summation takes place over all indices $\mathbf{m}\in\mathbb{Z}^d\cap[-C_1N_k/3, C_1N_k/3]^d$, with additional restrictions on $\mathbf{m}$ indicated below the relevant sum. We will estimate the integral representing $\mathscr{W}_1$ and the sum representing $\mathscr{W}_2$ separately.

Let us start with $\mathscr{W}_1$. There exists an absolute constant C_d such that

$$Q^*+3\mathbf{m}\delta_k \subseteq \big\{v : |u-v|\le C_d\delta_k\big\} \quad \text{for all} \quad \mathbf{m} \text{ with } |\mathbf{m}|\le 10d.$$

This means that the integral occurring in each summand of $\mathscr{W}_1$ can be bounded from above as follows,

$$\int_{Q^*+3\mathbf{m}\delta_k} \langle\lambda(u-v)\rangle^{-s}\,dv \le \int_{|u-v|\le C\delta_k} \langle\lambda(u-v)\rangle^{-s}\,dv. \tag{8.8}$$

The right hand side above is independent of $\mathbf{m}$, and the number of possible choices of $\mathbf{m}$ with $|\mathbf{m}|\le 10d$ is uniformly bounded. Thus squaring both sides of (8.8) above and summing in $\mathbf{m}$ yields

$$\mathscr{W}_1 \le C_d\Big[\int_{|u-v|\le C\delta_k} \langle\lambda(u-v)\rangle^{-s}\,dv\Big]^2.$$

After a change of variables $x=u-v$, we obtain

$$\begin{aligned}
\sqrt{\mathscr{W}_1} &\le \int_{|\mathbf{x}|\le C\delta_k} \langle\lambda x\rangle^{-s}\,d\mathbf{x}\\
&\le C_d\int_{r=0}^{C\delta_k} \langle\lambda r\rangle^{-s} r^{d-1}\,dr\\
&\le C_d\lambda^{-d}\int_0^{\lambda\delta_k} \langle t\rangle^{-s} t^{d-1}\,dt\\
&\le C_{d,s}\lambda^{-d}\times(\lambda\delta_k)^d\times\begin{Bmatrix} 1 & \text{if } \lambda\delta_k\le 1,\\ \Theta(\lambda\delta_k; s,d,;\log(1+\lambda\delta_k)) & \text{if } \lambda\delta_k>1\end{Bmatrix}\\
&\le C_{d,s}\delta_k^d\Theta(\zeta_k; s,d;\log(1+\zeta_k)).
\end{aligned} \tag{8.9}$$

In the second inequality of the sequence of steps above, we have made a polar change of coordinates $x = r\omega$, $r > 0$, $\omega \in \mathbb{S}^{d-1}$. This is followed by the scaling transformation $t = \lambda r$. The penultimate step invokes the estimate (4.3) derived in Lemma 4.1, with ϑ, ρ and τ in that lemma replaced by d, s and $\lambda\delta_k$ respectively. The last step combines the estimates arising in different regimes in a single closed form.

We now turn to $\mathscr{W}_2$. Since

$$C_d^{-1}|\mathbf{m}|\delta_k \le |x| \le C_d|\mathbf{m}|\delta_k \text{ for all } x \in \mathbf{m}\delta_k + [0,\delta_k]^d \text{ and all } \mathbf{m} \in \mathbb{Z}^d \text{ with } |\mathbf{m}| > 10d,$$

the sum represented by $\mathscr{W}_2$ may be interpreted as a lower Riemann sum. Replacing this sum by the corresponding integral yields

$$\begin{aligned}\mathscr{W}_2 &= \sideset{}{'}\sum_{|\mathbf{m}|>10d} \langle \lambda\delta_k\mathbf{m}\rangle^{-2s}\delta_k^{2d} \\ &\le C_{d,s}\delta_k^d \sideset{}{'}\sum_{|\mathbf{m}|>10d} \int_{\mathbf{m}\delta_k+[0,\delta_k]^d} \langle \lambda x\rangle^{-2s}\,dx \\ &\le C_{d,s}\delta_k^d \int_{\delta_k/C\le|x|\le C\delta_{k-1}} \langle \lambda x\rangle^{-2s}\,dx.\end{aligned}$$

The last inequality follows from the fact that the cubes $\mathbf{m}\delta_k + [0,\delta_k]^d$ are essentially disjoint, with

$$\bigcup\Big\{\mathbf{m}\delta_k + [0,\delta_k]^d : \mathbf{m} \in \mathbb{Z}^d \cap \big[-C_1N_k/3, C_1N_k/3\big]^d, |\mathbf{m}| \ge 10d\Big\} \subseteq \{x \in \mathbb{R}^d : \frac{\delta_k}{C} \le |x| \le C\delta_{k-1}\},$$

for a constant $C > \max(C_1 + \sqrt{d}, 1/(5d))$. The resulting integral is then estimated via the same sequence of spherical and scaling transformations as was used for $\mathscr{W}_1$:

$$\begin{aligned}\mathscr{W}_2 &\le C_{d,s}\delta_k^d \int_{\delta_k/C}^{C\delta_{k-1}} \langle \lambda r\rangle^{-2s} r^{d-1}\,dr \\ &\le C_{d,s}\delta_k^d\lambda^{-d} \int_{\lambda\delta_k/C}^{C\lambda\delta_{k-1}} t^{d-1}\langle t\rangle^{-2s}\,dt.\end{aligned}$$

While the univariate integral above is superficially similar to the the one considered in Lemma 4.1, it is important to note that the domain of integration here is an interval bounded strictly away from the origin, in contrast with Lemma 4.1 which deals with intervals whose left end point is the origin. As such, a direct application of Lemma 4.1 with $\vartheta = d$, $\rho = 2s$ and $\tau = C\lambda\delta_{k-1}$ would yield non-sharp upper bounds of the integral bounding $\mathscr{W}_2$, which could be significantly smaller in certain regimes of d and s. Accordingly, we estimate the last displayed integral directly on a case-by-case basis for different values of the integration limits $\lambda\delta_k, \lambda\delta_{k-1}$and the exponents d, s. This leads to

$$\mathscr{W}_2 \le C_{d,s}\widetilde{\mathscr{W}_2} \text{ where} \tag{8.10}$$

$$\widetilde{\mathscr{W}_2} := \delta_k^d \lambda^{-d} \times \left\{ \begin{array}{ll} (\lambda\delta_{k-1})^d & \text{if } \lambda\delta_{k-1} \le 1, \text{ any } s \text{ and } d, \\ (\lambda\delta_{k-1})^{d-2s} & \text{if } \lambda\delta_{k-1} > 1 \text{ and } d > 2s, \text{ any } \lambda\delta_k, \\ \log(\lambda\delta_{k-1}) - \log(\lambda\delta_k) & \text{if } \lambda\delta_{k-1} > \lambda\delta_k > 1 \text{ and } d = 2s, \\ (\lambda\delta_k)^{d-2s} & \text{if } \lambda\delta_{k-1} > \lambda\delta_k > 1 \text{ and } d < 2s, \\ \log(1+\lambda\delta_{k-1}) & \text{if } \lambda\delta_k \le 1 < \lambda\delta_{k-1} \text{ and } d = 2s, \\ 1 & \text{if } \lambda\delta_k \le 1 < \lambda\delta_{k-1} \text{ and } d < 2s. \end{array} \right\}. \tag{8.11}$$

We omit a detailed derivation of the estimates, opting to sketch the main idea instead. For the first four regimes, we observe that

$$\langle t \rangle^s = 1 + O((\lambda\delta_{k-1})^s) \text{ for any } s.$$

For the last two regimes, we ask the reader to decompose the domain of integration into two parts $[\lambda\delta_k/C, 1]$ and $[1, C\lambda\delta_{k-1}]$. The behaviour of the rational function integrand over the two subintervals then becomes easier to track.

Combining the estimates in (8.7), (8.9), (8.10) and (8.11) we obtain

$$\begin{aligned} \sum_{i_k} \big(w_k(\mathbf{i}_{k-1}, \bar{i}_k)\big)^2 &\le C_{d,s}(\mathscr{W}_1 + \mathscr{W}_2) \\ &\le C_{d,s}\Big(\big[\delta_k^d \Theta(\zeta_k; s, d; \log(1+\zeta_k))\big]^2 + \widetilde{\mathscr{W}_2} \Big) \le C_{d,s}\Xi_k, \end{aligned}$$

with Ξ_k as defined in (8.4). The last inequality follows from a direct comparison of the two summands in different regimes of $\lambda, \delta_k, \delta_{k-1}, s$ and d. We leave this to the interested reader. □

9. Estimation of $\mathscr{D}_k$ via a large deviation inequality

9.1. **A reduction of $\mathscr{D}_k$.** The estimates (8.2) and (8.5) prompt the following definitions. For $u \in [0,1]^d$, $1 \le r \le \ell + 1$ and with the convention that $\delta_0 = 1$, we set:

$$\mathbb{D}(r,\ell) = \mathbb{D}(r,\ell;u) := \left\{ \begin{array}{ll} \left\{ \mathbf{i}_\ell \in \mathbb{I}(\ell,d) : \delta_r\sqrt{d} < \Delta(u,\mathbf{i}_\ell) \le \delta_{r-1}\sqrt{d} \right\} & \text{if } 1 \le r \le \ell, \\ \left\{ \mathbf{i}_\ell \in \mathbb{I}(\ell,d) : \Delta(u,\mathbf{i}_\ell) \le \delta_\ell\sqrt{d} \right\} & \text{if } r = \ell+1. \end{array} \right\} \tag{9.1}$$

It is clear that for any $u \in [0,1]^d$,

$$\mathbb{I}(\ell,d) = \bigsqcup_{r=1}^{\ell+1} \mathbb{D}(r,\ell;u). \tag{9.2}$$

We also define

$$Z_r(u;\ell) := \sum_{\mathbf{i}_\ell \in \mathbb{D}(r,\ell)} X_\ell(\mathbf{i}_\ell). \tag{9.3}$$

The relevance of the quantities $\mathbb{D}(r,\ell)$ and $Z_r(u;\ell)$ is made clearer in the next lemma.

Lemma 9.1. *With $\mathbb{D}(r,\ell)$ and $Z_r(u;\ell)$ as above and ζ_k as in Lemma 8.2, the following estimate holds:*

$$|\mathscr{D}_k(u,\lambda;s)| \le C_{d,s}\Big[\delta_k^d\delta_{k-1}^d\sum_{r=1}^{k-1} Z_r(u;k-1)\langle\lambda\delta_r\rangle^{-2s} + Z_k(u,k-1)\Xi_k\Big], \tag{9.4}$$

where Ξ_k is as in (8.4).

Proof. We use the estimates obtained in Lemmas 8.1 and 8.2 to bound $\mathscr{D}_k$ as in (6.1) :

$$\begin{aligned}
\mathscr{D}_k &= \sum_{\mathbf{i}_{k-1}\in\mathbb{I}(k-1,d)} X_{k-1}(\mathbf{i}_{k-1})\sum_{\bar{i}_k}(w_k(\bar{i}_k,\mathbf{i}_{k-1}))^2\\
&= \sum_{r=1}^{k}\sum_{\mathbf{i}_{k-1}\in\mathbb{D}(r,k-1;u)} X_{k-1}(\mathbf{i}_{k-1})\sum_{\bar{i}_k}(w_k(\bar{i}_k,\mathbf{i}_{k-1}))^2 \quad \text{by (9.2) with } \ell=k-1,\\
&\le C_{d,s}\Bigg[\delta_k^d\delta_{k-1}^d\sum_{r=1}^{k-1}\sum_{\mathbf{i}_{k-1}\in\mathbb{D}(r,k-1;u)} X_{k-1}(\mathbf{i}_{k-1})\langle\lambda\Delta(u,\mathbf{i}_{k-1})\rangle^{-2s}\\
&\qquad\qquad + \sum_{\mathbf{i}_{k-1}\in\mathbb{D}(k,k-1;u)} X_{k-1}(\mathbf{i}_{k-1})\sum_{\bar{i}_k}(w_k(\bar{i}_k,\mathbf{i}_{k-1}))^2\Bigg] \quad \text{by (8.2)},\\
&\le C_{d,s}\Bigg[\delta_k^d\delta_{k-1}^d\sum_{r=1}^{k-1}\langle\lambda\delta_r\sqrt{d}\rangle^{-2s}\sum_{\mathbf{i}_{k-1}\in\mathbb{D}(r,k-1;u)} X_{k-1}(\mathbf{i}_{k-1}) +\\
&\qquad\qquad \Xi_k\sum_{\mathbf{i}_{k-1}\in\mathbb{D}(k,k-1;u)} X_{k-1}(\mathbf{i}_{k-1})\Bigg] \quad \text{by (8.5)}.
\end{aligned}$$

The penultimate inequality in the sequence of steps above follows from (9.2), with $\ell=k-1$. The last expression is essentially the right hand side of (9.4), in view of (9.3). □

Lemma 9.2. *For $\mathbb{P}^*$-a.e. $\omega\in\Omega$, there exists a constant $C=C_\omega>0$ such that for all $u\in[0,1]^d$ and all indices r,k with $1\le r\le k-1$, the quantity $Z_r(u,\ell)$ defined as in* (9.3) *obeys the following bound:*

$$Z_r(u,k-1) \le C(\delta_{r-1}/\delta_{k-1})^d\prod_{m=r}^{k-1} p_m. \tag{9.5}$$

Proof. Let us set $\ell=k-1$ and recall the definition (9.1) of $\mathbb{D}(r,\ell)$. Since $\mathbb{D}(r,\ell)$ only involves cubes whose distance from u is at most $\sqrt{d}\delta_{r-1}$, we observe that there is a constant C_d depending only on the dimension d such that $\sqrt{d}\delta_{r-1}$-neighbourhood of u can be covered by at most C_d cubes from the $(r-1)$th generation, each of sidelength δ_{r-1}. In other words, given any $u\in[0,1]^d$, there is a collection $\mathbb{I}_u(r-1,d)\subseteq\mathbb{I}(r-1,d)$ of cardinality at most C_d with the property that every $\mathbf{i}_\ell\in\mathbb{D}(r,\ell)$ is of the form $\mathbf{i}_\ell=(\mathbf{i}_{r-1},\mathbf{j})$ for some $\mathbf{i}_{r-1}\in\mathbb{I}_u(r-1,d)$ and some multi-index $\mathbf{j}$ of length $\ell-r+1$. This means that $Z_r(u,\ell)$ can be no more than the total number of basic cubes of the ℓ-th generation of the Cantor set that are descended from the cubes $Q(\mathbf{i}_{r-1})$, $\mathbf{i}_{r-1}\in\mathbb{I}_u(r-1,d)$. In the notation of Section 12.2, this means that

$$Z_r(u,\ell)\le C_d\sup\left\{q_\ell[\mathbf{i}_{r-1}] : \mathbf{i}_{r-1}\in\mathbb{I}(r-1,d)\right\}.$$

According to Lemma 12.1, this number is bounded by the right hand side of (9.5). □

9.2. **Proof of Proposition 6.2.**

Proof. Let us recall the description (5.4) of the quantity Ψ_k occurring in the right hand side of (6.6). We aim to show that $\mathscr{D}_k$ is bounded above by this quantity.

Since $Z_k(u, k-1)$ is bounded above by a constant C_d uniformly in u, the estimate (9.4) combined with (9.5) yields

$$\mathscr{D}_k \leq C_{d,s}\Big[\delta_k^d \delta_{k-1}^d \sum_{r=1}^{k-1} \langle \lambda\delta_r\rangle^{-2s} \big(\delta_{r-1}/\delta_{k-1}\big)^d \prod_{m=r}^{k-1} p_m + \Xi_k\Big]. \tag{9.6}$$

We recall from (8.4) that Ξ_k is a function of $\zeta_k = \max(\lambda\delta_k, 1)$ and ζ_{k-1}. The presence of the terms $\langle\lambda\delta_r\rangle$ and ζ_k and ζ_{k-1} in the sum above suggests that we should estimate it in two separate cases depending on the relative sizes of λ and δ_k and δ_{k-1}.

Case 1: $\lambda\delta_k > 1$. In this case $\lambda\delta_r$ is larger than 1 for all $1 \leq r \leq k-1$, as a result of which $\langle\lambda\delta_r\rangle$ is comparable to $\lambda\delta_r$. In light of this, the bound (9.6) reduces to

$$\mathscr{D}_k \leq C_d \delta_k^d \lambda^{-2s} \prod_{m=1}^{k-1} p_m \Big[\sum_{r=1}^{k-1} \delta_r^{-2s} \delta_{r-1}^d \prod_{m=1}^{r-1} p_m^{-1}\Big] + \Xi_k. \tag{9.7}$$

Using the definition of δ_r that follows from our choice of parameters (2.20) and (2.21), we verify that there exists a constant $C_{d,s}$ such that

$$\sum_{r=1}^{k-1} \delta_r^{-2s} \delta_{r-1}^d \prod_{m=1}^{r-1} p_m^{-1} \leq C_{d,s} \times \mathscr{D}_k', \text{ where} \tag{9.8}$$

$$\mathscr{D}_k' := \begin{Bmatrix} 1 & \text{if } s < \varrho/2, \\ \delta_{k-1}^{\varrho-2s} N_{k-1}^B & \text{if } s \geq \varrho/2, \end{Bmatrix}, \quad \text{with} \quad B = B(d,s,\varepsilon) = \begin{cases} 2s & \text{if } \varepsilon > 0, \\ 2s+\gamma & \text{if } \varepsilon = 0. \end{cases} \tag{9.9}$$

The constant γ here is the same one that appears in (2.20). Inserting the estimate (9.8) as given by (9.9) into the expression in (9.7), and comparing it with the size of Ξ_k as given by (8.4), we arrive at the following bound on $\mathscr{D}_k$:

$$\mathscr{D}_k \leq C_{d,s} \times \begin{Bmatrix} \lambda^{-2s}\delta_k^d \prod_{m=1}^{k-1} p_m & \text{if } s < \dfrac{\varrho}{2}, \\ \lambda^{-2s}\delta_k^d \delta_{k-1}^{\varrho-2s} N_{k-1}^B \prod_{m=1}^{k-1} p_m & \text{if } \dfrac{\varrho}{2} \leq s < d, \\ \lambda^{-2d}\Phi(\lambda) & \text{if } s \geq d, \end{Bmatrix} \text{ for } \lambda\delta_k > 1. \tag{9.10}$$

Let us recall that Φ is a slowly growing function given by (5.6). We ask the reader to verify that the right hand side above agrees with the expression for Ψ_k as specified by (5.4), with the constant C in (5.5) being at least B.

Case 2: $\lambda\delta_k \leq 1 < \lambda\delta_{k-1}$. In this case, we still have that $\zeta_r = \lambda\delta_r \geq 1$ for $1 \leq r \leq k-1$, as a result of which (9.7) continues to hold, with the first summand obeying the same bound

given by (9.8) and (9.9). The only distinction is that Ξ_k is now given by

$$\Xi_k = \delta_k^d \delta_{k-1}^d \Theta(\lambda\delta_{k-1}; 2s, d; \log(1+\lambda\delta_{k-1})),$$

as can be seen from (8.4). A case-by-case comparison as before now leads to

$$\mathscr{D}_k \le C_{d,s} \times \left\{ \begin{array}{ll} \lambda^{-2s}\delta_k^d \prod_{m=1}^{k-1} p_m & \text{if } s < \frac{\varrho}{2}, \\ \lambda^{-2s}\delta_k^d \delta_{k-1}^{\varrho-2s} N_{k-1}^B \prod_{m=1}^{k-1} p_m & \text{if } \frac{\varrho}{2} \le s < \frac{d}{2}, \\ \lambda^{-d}\delta_k^d N_{k-1}^B & \text{if } s \ge \frac{d}{2}. \end{array} \right\} \text{ for } \lambda\delta_k \le 1 < \lambda\delta_{k-1}. \tag{9.11}$$

This too coincides with Ψ_k as given by (5.4). Note that the range $\varrho/2 \le s < d/2$ is nonempty only if $\varepsilon > 0$.

Case 3: $\lambda\delta_{k-1} < 1$, which is equivalent to $\zeta_{k-1} = 1$. The new feature of this case is that $\lambda\delta_r$ is larger than 1 for small values of r and less than 1 otherwise, as a result of which $\langle\lambda\delta_r\rangle$ is comparable respectively to $\lambda\delta_r$ and a constant in these two regimes. Furthermore, the last summand Ξ_k in (9.6) is $\delta_k^d\delta_{k-1}^d$, as can be seen from (8.4). Accordingly, we split the sum in (9.6) in three parts:

$$\mathscr{D}_k \le \mathscr{D}_k^* + \mathscr{D}_k^{**} + C_d\delta_k^d\delta_{k-1}^d, \text{ where} \tag{9.12}$$

$$\mathscr{D}_k^* \le C_d\big[\lambda^{-2s}\delta_k^d \prod_{m=1}^{k-1} p_m\big] \times \Big[\sum\nolimits^* \delta_r^{-2s}\delta_{r-1}^d \prod_{m=1}^{r-1} p_m^{-1}\Big], \tag{9.13}$$

$$\mathscr{D}_k^{**} \le C_d\big[\delta_k^d \prod_{m=1}^{k-1} p_m\big] \times \Big[\sum_r^{**} \delta_{r-1}^d \prod_{m=1}^{r-1} p_m^{-1}\Big]. \tag{9.14}$$

Here $\sum^*$ denotes the sum over all indices $1 \le r \le k-1$ such that $\lambda\delta_r \ge 1$, and $\sum^{**}$ denotes the sum over the complementary set of indices r. The sum appearing in $\mathscr{D}_k^*$ is very similar to $\mathscr{D}_k'$ appearing in Case 1. The estimation of $\mathscr{D}_k^*$ therefore proceeds in an essentially identical manner, and yields

$$\mathscr{D}_k^* \le \lambda^{-2s}\delta_k^d \prod_{m=1}^{k-1} p_m \times \left\{ \begin{array}{ll} 1 & \text{if } s < \varrho/2, \\ \lambda^{2s-\varrho}\Phi(\log\lambda) & \text{if } s \ge \varrho/2. \end{array} \right\}. \tag{9.15}$$

The second sum $\mathscr{D}_k^{**}$ is estimated as follows,

$$\begin{aligned} \mathscr{D}_k^{**} &\le C_{d,s}\big[\delta_k^d \prod_{m=1}^{k-1} p_m\big] \times \left\{ \begin{array}{ll} \sum^{**} \delta_r^{d(1-\varepsilon)} N_r^{d(1-\varepsilon)} & \text{if } \varepsilon > 0 \\ \sum^{**} \delta_r^d N_r^{d(1+\gamma)} & \text{if } \varepsilon = 0. \end{array} \right\} \\ &\le C_{d,s}\big[\delta_k^d \prod_{m=1}^{k-1} p_m\big] \times \Phi(\lambda)\lambda^{-\varrho}. \end{aligned} \tag{9.16}$$

Comparing the estimates in (9.15) and (9.16) and inserting them into (9.12) we find that $\mathscr{D}_k^{**} + \delta_k^d \delta_{k-1}^d \leq \mathscr{D}_k^*$, thus proving $\mathscr{D}_k \leq C_{d,s}\mathscr{D}_k^*$. Since $\mathscr{D}_k^*$ equals Ψ_k in the regime $\lambda\delta_{k-1} < 1$, we are done. □

10. Sharpness: Proof of Theorem 1.3 (c)

Let us recall the random Cantor construction laid out in Section 2.4. Given a Riemannian manifold (M, g) and a submanifold Σ of dimension d, we obtained a measure space $(\Omega, \mathbb{P}^*)$ such that $\mathbb{P}^*$-almost every point $\omega \in \Omega$ generated a Cantor-like set $E = E(\omega) \subseteq \Sigma$ of Hausdorff dimension $d(1-\varepsilon)$, equipped with a natural measure $\mu = \mu(\omega)$. Our main objective in this section is to prove the almost sure lower bound (1.10) for $M = \mathbb{S}^n$.

The overall structure of the proof is very similar to [8]. However, for the sake of completeness, we include it in its entirety, since at critical junctures of the argument, well-known properties of the Lebesgue measure have to be replaced by their analogues for μ. The important tools in the proof are the following.

10.1. **Summary of results.**

Proposition 10.1 (Lower bound for rough spectral projectors)**.** *For any Riemannian manifold (M, g),*

$$\limsup_{\substack{m\to\infty \\ m\in\mathbb{N}}} m^{-\delta_p}\Phi(m)||\mathbf{1}_{m/2}||_{L^2(M)\to L^p(\nu)} > 0, \tag{10.1}$$

where $\mathbf{1}_\lambda$ denotes the rough spectral projector $\mathbf{1}_\lambda = \mathbf{1}_{\sqrt{-\Delta}\in[\lambda,\lambda+\frac{1}{2})}$.

In other words, there exists a countably infinite increasing sequence of spectral parameters $\{\theta_k : k \geq 1\}$ (not necessarily eigenvalues) with the following properties.

(a) Each θ_k is a non-negative half-integer such that $[\theta_k, \theta_k + 1/2) \cap Spec(-\Delta_g) \neq \emptyset$.

(b) For $\mathbb{P}^*$-almost every $\omega \in \Omega$ and every exponent $p \geq p_0$, one can find a constant $C = C(\omega, p) > 0$ satisfying

$$||\mathbf{1}_{\theta_k}||_{L^2(M)\to L^p(\nu)} \geq C\theta_k^{\delta_p}\Phi(\theta_k)^{-1} \quad \text{for all } k \geq 1. \tag{10.2}$$

Proposition 10.2 (Lower bound for smooth spectral projectors)**.** *For any Riemannian manifold (M, g),*

$$\limsup_{\lambda\to\infty} \lambda^{-\delta_p}\Phi(\lambda)||\mathcal{T}_\lambda||_{L^2(M)\to L^p(\nu)} > 0, \tag{10.3}$$

where $\mathscr{T}_\lambda = \chi(\sqrt{-\Delta_g} - \lambda)$ is the smooth projection operator defined in (3.2).

Proposition 10.3 (A ball condition for random measures)**.** *For $\mathbb{P}^*$-almost every $\omega \in \Omega$, the following statements hold for the random sets $E_\omega \subseteq [0, 1]^d$ constructed in Section 2.4.*

(a) There exists a constant $C = C_\omega > 0$ such that

$$\sup\Big\{\mu_\omega\big(B(v_0; r)\big)r^{-d(1-\varepsilon)}\big[\Phi(1/r)\big]^{-1} : r > 0, v_0 \in E_\omega\Big\} \leq C_\omega < \infty. \tag{10.4}$$

(b) For every $v_0 \in E(\omega)$, there exists a constant $C_0 = C(v_0; \omega) > 0$ such that for all $r > 0$,

$$\inf\Big\{\mu_\omega\big(B(v_0; r)\big)r^{-d(1-\varepsilon)}\big[\Phi(1/r)\big]^{-1} : r > 0\Big\} \geq C_0^{-1}. \tag{10.5}$$

10.2. **Proof of Theorem 1.3 (c), assuming Proposition 10.1.**

Proof. For $M = \mathbb{S}^n$, we have explicit information about the location of the eigenvalues of the Laplace-Beltrami operator:

$$\mathrm{Spec}(-\Delta_{\mathbb{S}^n}) = \big\{\lambda_\ell^2 = \ell(\ell+n-1) : \ell \in \mathbb{N}\big\}.$$

The gap between $\lambda_{\ell+1}$ and λ_{λ_ℓ} approaches 1 as $\ell \to \infty$. Thus for any large integer $m \in \mathbb{N}$, the interval $[m/2, m/2+1/2)$ contains at most one value of λ_ℓ. Thus in this context, Proposition 10.1 provides an increasing sequence of positive half-integers θ_k such that each interval $[\theta_k, \theta_k + 1/2)$ contains the square root of exactly one eigenvalue, say λ_{ℓ_k}. Since $\mathbf{1}_{\theta_k}\varphi_{\ell_k} = \varphi_{\ell_k}$ for any eigenfunction φ_{ℓ_k} associated with the eigenvalue λ_{ℓ_k}, the desired conclusion (1.10) follows from (10.2) for the subsequence φ_{ℓ_k} of L^2-normalized spherical harmonics. □

10.3. **Proof of Proposition 10.1 assuming Proposition 10.2.**

Proof. We prove this by contradiction. If the limit superior in (10.1) is zero, then we can find $\kappa_m \to 0$ such that

$$||\mathbf{1}_{\frac{m}{2}}||_{L^2(M)\to L^p(\nu)} \le \kappa_m \Phi(m)^{-1} m^{\delta_p} \quad \text{for all } m. \tag{10.6}$$

Without loss of generality and after choosing a slower decaying function if necessary, we may assume that $\lambda \mapsto \kappa_\lambda$ is a continuous function on $\mathbb{R}_{\ge 0}$ decreasing to zero at infinity, such that $f(\lambda) = \kappa_\lambda \lambda^{\delta_p}\Phi(\lambda)^{-1}$ obeys the following properties:

- f is increasing in λ for sufficiently large λ,
- f satisfies a doubling condition, i.e., there exists a constant C such that

$$f(2\lambda) \le C f(\lambda) \text{ for all large } \lambda. \tag{10.7}$$

The hypotheses above are applied towards estimating the operator norm of $\mathcal{T}_\lambda$, in the following way. We observe that

$$\mathrm{Id} = \sum_{m\in\mathbb{Z}} \mathbf{1}_{m/2} = \sum_{m\in\mathbb{Z}} \mathbf{1}_{m/2} \circ \mathbf{1}_{m/2}, \quad \text{which implies}$$

$$\mathscr{T}_\lambda = \sum_{m\in\mathbb{Z}} \mathbf{1}_{m/2} \circ \mathbf{1}_{m/2} \circ \mathscr{T}_\lambda. \tag{10.8}$$

Taking the operator norm of both sides of (10.8) and invoking (10.6), we arrive at the estimate

$$\begin{aligned}
||\mathcal{T}_\lambda||_{L^2(M)\to L^p(\nu)} &\le \sum_{m\in\mathbb{Z}} ||\mathbf{1}_{m/2} \circ \mathbf{1}_{m/2} \circ \mathcal{T}_\lambda||_{L^2(M)\to L^p(\nu)} \\
&\le \sum_{m\in\mathbb{Z}} ||\mathbf{1}_{m/2}||_{L^2(M)\to L^p(\nu)} \times ||\mathbf{1}_{m/2} \circ \mathscr{T}_\lambda||_{L^2(M)\to L^2(M)} \\
&\le \sum_{m\in\mathbb{Z}} \kappa_m \Phi(m)^{-1} m^{\delta_p} \sup_{j:\lambda_j\in J_m} |\chi(\lambda_j - \lambda)| \\
&\le C_N \sum_{m\in\mathbb{Z}} \kappa_m \Phi(m)^{-1} m^{\delta_p} \big(1 + \mathrm{dist}(\lambda, J_m)\big)^{-N}. \qquad (10.9)
\end{aligned}$$

Here J_m denotes the interval $[m/2, (m+1)/2)$. The third inequality in the sequence above follows from (10.6) and the spectral theorem. The fourth inequality, which holds for any

positive integer N, uses the fact that χ is Schwartz. We now proceed to estimate the sum in (10.9) in three parts, depending on the relative position of the running index m with respect to m^*, where m^* denotes the unique integer such that $\lambda \in J_{m^*}$. Thus $m^* \leq 2\lambda$.

For $m \leq m^* - 2$ we set $r = m^* - m$, so that $\operatorname{dist}(\lambda, J_m) \geq \operatorname{dist}(J_m, J_{m^*}) = (r-1)/2$. Using the fact that $f(\lambda) = \kappa_\lambda \lambda^{\delta_p} \Phi(\lambda)^{-1}$ increases with λ, we obtain

$$\begin{aligned}\sum_{m \leq m^*-2} \kappa_m \Phi(m)^{-1} m^{\delta_p} \big(1 + \operatorname{dist}(\lambda, J_m)\big)^{-N} &\leq f(m^*) \sum_{r \geq 2} ((r-1)/2)^{-N} \\ &\leq C f(2\lambda) \leq C f(\lambda) = C \kappa_\lambda \Phi(\lambda)^{-1} \lambda^{\delta_p}, \end{aligned} \tag{10.10}$$

where the last inequality follows from the doubling property (10.7). For m in the range $|m - m^*| \leq 1$, the estimate

$$\sum_{m=m^*-1}^{m^*+1} \kappa_m \Phi(m)^{-1} m^{\delta_p} \big(1 + \operatorname{dist}(\lambda, J_m)\big)^{-N} \leq C \kappa_\lambda \Phi(\lambda)^{-1} \lambda^{\delta_p} \tag{10.11}$$

is easy to verify from the properties of f; in fact, each one of the three summands is comparable to the right hand side, by the doubling property (10.7). For $m \geq m^* + 2$, we set $r = m - m^*$, so that once again we have $\operatorname{dist}(\lambda, J_m) \geq (r-1)/2$. Since κ_m and $1/\Phi(m)$ both decrease with m, this leads to

$$\begin{aligned}\sum_{m=m^*+1}^{\infty} \kappa_m \Phi(m)^{-1} m^{\delta_p} \big(1 + \operatorname{dist}(\lambda, J_m)\big)^{-N} &\leq \kappa_{m^*} \sum_{r=2}^{\infty} \Phi(m^* + r)^{-1} (m^* + r)^{\delta_p} ((r-1)/2)^{-N} \\ &\leq \kappa_{m^*} \Phi(m^*)^{-1} \sum_{r=2}^{\infty} (m^* + r)^{\delta_p} ((r-1)/2)^{-N} \\ &\leq C_N \kappa_{m^*} \Phi(m^*)^{-1} \sum_{r=2}^{\infty} (m^*)^{\delta_p} r^{\delta_p - N} \\ &\leq C f(m^*) \leq C f(\lambda) = C \kappa_\lambda \Phi(\lambda)^{-1} \lambda^{\delta_p}. \end{aligned} \tag{10.12}$$

The fourth inequality holds provided one chooses $N > \delta_p + 1$. The subsequent inequality is justified exactly as in (10.10).

Combining the estimates obtained in (10.9)-(10.12), we find that

$$||\mathcal{T}_\lambda||_{L^2(M) \to L^p(\nu)} \leq C \kappa_\lambda \lambda^{\delta_p} \text{ for all sufficiently large } \lambda.$$

Since $\kappa_\lambda \to 0$, this implies that

$$\limsup_{\lambda \to \infty} \lambda^{-\delta_p} \Phi(\lambda) ||\mathcal{T}_\lambda||_{L^2(M) \to L^p(\nu)} = 0,$$

contradicting the conclusion of Proposition 10.2. □

10.4. **Proof of Proposition 10.2, assuming Proposition 10.3.**

Proof. By Theorem 3.1, $\mathscr{T}_\lambda$ admits the representation given by (3.6). As before, in view of the fast decay

$$||\mathscr{R}_\lambda||_{L^2(M) \to L^p(\nu)} \leq C_N \lambda^{-N} \text{ for any } N \geq 1 \text{ and any } p \geq 2,$$

we continue to denote by $\mathscr{T}_\lambda$ the leading term in (3.6). Proposition 10.2 demands the existence of a constant $c > 0$ such that

$$||\mathcal{T}_\lambda||^2_{L^2(M)\to L^p(\nu)} = ||\mathcal{T}_\lambda\mathcal{T}_\lambda^*||_{L^{p'}(\nu)\to L^p(\nu)} \geq c\Phi(\lambda)^{-1}\lambda^{2\delta_p} \quad \text{for all large } \lambda, \tag{10.13}$$

where $\mathcal{T}_\lambda$ denotes the restriction of the operator $\mathscr{T}_\lambda$ as given by (3.1). From the discussion in Section 3, we find that $\mathcal{T}_\lambda\mathcal{T}_\lambda^*$ is an integral operator of the form (3.7), with integration kernel $\mathscr{K}_\lambda$ given by

$$\mathscr{K}_\lambda(x(u), x(v)) = \lambda^{n-1}\int \exp\big[-i\lambda\big(d_g(x(u), y) - d_g(x(v), y)\big)\big]a\big(x(u), y\big)a\big(x(v), y\big)\, dy.$$

We have also recorded in Section 3 (specifically in Theorem 3.1(a)) that a_λ is a smooth function that does not vanish for $(x, y) \in S$ with $d_g(x, y) \in [c_3\kappa, c_4\kappa]$. Here $\kappa > 0$ is a fixed small constant, and S has been defined in (3.5).

For $\mathbb{P}^*$-almost every $\omega \in \Omega$ obeying the conclusion of Proposition 10.3, let us fix $v_0 \in E_\omega$. We are thus ensured of the two estimates (10.4) and (10.5). For all sufficiently large λ, we choose a test function f_λ such that

$$f_\lambda(x(v)) = \lambda^{d(1-\varepsilon)/p'}\psi(\lambda(v - v_0)). \tag{10.14}$$

Here ψ is a smooth non-negative function supported in a small ball of radius σ (a small absolute constant soon to be specified) centred at the origin and identically one on a concentric ball of half the radius. By choosing a smaller σ if necessary, we can ensure that the function $(v, y) \mapsto a(x(v), y)$ does not change sign and stays bounded away from zero for $(x(v), y) \in S$ and $x(v) \in \operatorname{supp}(f_\lambda)$. It follows from (2.24) and (10.4) that

$$\begin{aligned} ||f_\lambda||^{p'}_{L^{p'}(\nu)} &= \int |f_\lambda(x(v))|^{p'}\sqrt{\det(g(v))}\, d\mu(v) \\ &\leq C_\sigma\lambda^{d(1-\varepsilon)}\mu\big(B(v_0; \sigma\lambda^{-1})\big) \leq C_\sigma\Phi(\lambda). \end{aligned} \tag{10.15}$$

An additional relevance of σ is that

$$|d_g(x(u), y) - d_g(x(v), y)| \leq d_g(x(u), x(v)) \leq C\sigma\lambda^{-1} \quad \text{for } x(u), x(v) \in \operatorname{supp}(f_\lambda),$$

so for σ small enough, we can ensure that

$$\begin{aligned} &\operatorname{Re}\Big[\mathscr{K}_\lambda(x(u), x(v))\Big] \\ &\quad\geq \lambda^{n-1}\int \operatorname{Re}\Big[\exp\big\{-i\lambda\big(d_g(x(u), y) - d_g(x(v), y)\big)\big\}\Big]a\big(x(u), y\big)a\big(x(v), y\big)\, dy \\ &\quad\geq \cos(C\sigma)\lambda^{n-1}\int a\big(x(u), y\big)a\big(x(v), y\big)\, dy \geq c_0\lambda^{n-1} \end{aligned}$$

for some fixed $c_0 > 0$. Thus for $x(u) \in \operatorname{supp}(f_\lambda)$, with f_λ as in (10.14),

$$\begin{aligned} \big|\mathcal{T}_\lambda\mathcal{T}_\lambda^* f_\lambda(x(u))\big| &\geq c_0\lambda^{n-1}\int f_\lambda(x(v))\sqrt{\det(g(v))}\, d\mu(v) \\ &\geq C^{-1}\lambda^{n-1+\frac{d(1-\varepsilon)}{p'}}\mu\big[B(v_0; \sigma/(2\lambda))\big] \\ &\geq C_\sigma^{-1}\lambda^{n-1+\frac{d(1-\varepsilon)}{p'}}\lambda^{-d(1-\varepsilon)}\Phi(\lambda)^{-1} = C_\sigma^{-1}\lambda^{n-1-\frac{d(1-\varepsilon)}{p}}\Phi(\lambda)^{-1}, \end{aligned}$$

where the last inequality is a consequence of (10.5). The pointwise bound above implies that

$$\begin{aligned} ||\mathcal{T}_\lambda \mathcal{T}_\lambda^* f_\lambda||_{L^p(\nu)}^p &\geq C_\sigma^{-1} \lambda^{(n-1)p-d(1-\varepsilon)} \Phi(\lambda)^{-1} \nu\big(\mathrm{supp}(f_\lambda)\big) \\ &\geq C_\sigma^{-1} \lambda^{(n-1)p-d(1-\varepsilon)} \Phi(\lambda)^{-1} \mu\big[B(v_0; \sigma\lambda^{-1}/2)\big] \\ &\geq C_\sigma^{-1} \lambda^{(n-1)p-2d(1-\varepsilon)} \Phi(\lambda)^{-1} = C_\sigma^{-1} \Phi(\lambda)^{-1} \lambda^{2p\delta_p}, \end{aligned} \tag{10.16}$$

where we have used (10.5) again in the last inequality. Combining (10.15) with (10.16) leads us to the desired conclusion (10.13). □

10.5. **Proof of Proposition 10.3.**

Proof. Fix any $v_0 \in E$ and any $0 < r < 1$. Since

$$\mu\big[B(v_0; r\big] = \lim_{k\to\infty} \mu_k\big[B(v_0; r)\big],$$

the desired inequality (10.4) follows from a similar inequality with μ_k replacing μ. It follows from (2.8) that

$$\begin{aligned} \mu_k\big[B(v_0; r)\big] &= \frac{1}{P_k \delta_k^d} \sum_{\mathbf{i}_k} X_k(\mathbf{i}_k) \int_{B(v_0;r)} 1_{Q(\mathbf{i}_k)}(x)\, dx \\ &= \frac{1}{P_k \delta_k^d} \sum_{\mathbf{i}_k} X_k(\mathbf{i}_k) \big|B(v_0; r) \cap Q(\mathbf{i}_k)\big|. \end{aligned} \tag{10.17}$$

To establish (10.4), let us choose the unique scale ℓ^* such that

$$\delta_{\ell^*+1} < r \leq \delta_{\ell^*}. \tag{10.18}$$

The relevance of ℓ^* is that $B(v_0; r)$ can be covered by cubes of the form $\{Q(\mathbf{i}_{\ell^*}) : \mathbf{i}_{\ell^*} \in \mathbb{I}^*[v_0; r]\}$, where the cardinality of $\mathbb{I}^*[v_0; r]$ is at most a constant C_d depending only on d. Thus the cubes $Q(\mathbf{i}_k)$ that contribute to the sum in (10.17) are those descended from $Q(\mathbf{i}_{\ell^*})$ for some $\mathbf{i}_{\ell^*} \in \mathbb{I}^*[v_0; r]$. In other words, the sum ranges over multi-indices $\mathbf{i}_k \in \mathbb{I}(k, d)$ whose projection onto the first ℓ^* coordinates yields some $\mathbf{i}_{\ell^*} \in \mathbb{I}^*[v_0; r]$. This leads to the following estimate:

$$\begin{aligned} \mu_k\big[B(v_0; r)\big] &\leq \frac{1}{P_k \delta_k^d} \sum_{\mathbf{i}_{\ell^*} \in \mathbb{I}^*[v_0;r]} \sum\nolimits' X_k(\mathbf{i}_k) \delta_k^d \\ &= \frac{1}{P_k} \sum_{\mathbf{i}_{\ell^*} \in \mathbb{I}^*[v_0;r]} q_k[\mathbf{i}_{\ell^*}] \\ &\leq \frac{C_d}{P_k} \sup\Big\{ q_k[\mathbf{i}_{\ell^*}] : \mathbf{i}_{\ell^*} \in \mathbb{I}(\ell^*, d) \Big\} \\ &\leq \frac{C_d}{P_k} \left(\frac{\delta_{\ell^*}}{\delta_k}\right)^d \prod_{m=\ell^*+1}^{k} p_m. \end{aligned} \tag{10.19}$$

The sum $\sum'$ in the first displayed line above ranges over all multi-indices $\mathbf{i}_k$ which project onto some $\mathbf{i}_{\ell^*} \in \mathbb{I}^*[v_0; r]$ in the first ℓ^* coordinates. Thus the quantity $q_k[\mathbf{i}_{\ell^*}]$ that appears in the second line is the same as the one defined in (12.5) in Section 12.2, namely the number of basic cubes of the kth generation descended from $Q(\mathbf{i}_{\ell^*})$. Lemma 12.1 then provides the

upper bound in (10.19). A simplification of this last term using (2.13), (2.20) and (2.21) yields

$$\frac{C_d}{P_k}\left(\frac{\delta_{\ell^*}}{\delta_k}\right)^d \prod_{m=\ell^*+1}^{k} p_m = \begin{cases} \delta_{\ell^*}^{d(1-\varepsilon)} & \text{if } \varepsilon > 0, \\ \delta_{\ell^*}^{d} N_{\ell^*}^{d\gamma} & \text{if } \varepsilon = 0. \end{cases}$$

In view of (10.18), both expressions above are dominated by $r^{d(1-\varepsilon)}\Phi(1/r)$, completing the proof of (10.4).

We turn to the inequality (10.5). For any $v_0 \in \mathbb{R}^d$, the ball $B(v_0; r)$ contains a cube $Q = Q(v_0; r)$ centred at v_0 and of sidelength $2r/\sqrt{d}$. It follows from (10.18) that there exists a multi-index $\widehat{\mathbf{i}}_{\ell^*+2} \in \mathbb{I}(\ell^*+2, d)$ such that

$$v_0 \in \widehat{Q} = Q(\widehat{\mathbf{i}}_{\ell^*+2}) \subset Q,$$

provided the parameter N in (2.20) and (2.21) is chosen large enough relative to d. The relation (10.17) then leads to the following lower bound: for $k \geq \ell^* + 3$,

$$\mu_k[B(v_0; r)] \geq \frac{1}{P_k \delta_k^d} \hat{\sum_{\mathbf{i}_k}} X_k(\mathbf{i}_k)\delta_k^d = \frac{1}{P_k} q_k[\widehat{\mathbf{i}}_{\ell^*+2}], \tag{10.20}$$

where $\hat{\sum}$ indicates summation over all multi-indices $\mathbf{i}_k$ whose projection onto the first ℓ^*+2 coordinates yields $\widehat{\mathbf{i}}_{\ell^*+2}$. In other words, we only restrict attention to cubes $Q(\mathbf{i}_k)$ that are descended from $\widehat{Q}$. If additionally we assume that $v_0 \in E$, then the right hand side of (10.20) is guaranteed to be nonzero; in fact, by Lemma 12.2 we can estimate it from below by

$$\mu_k[B(v_0; r)] \geq \frac{1}{P_k} q_k[\widehat{\mathbf{i}}_{\ell^*+2}] \geq \frac{1}{CP_k}\left(\frac{\delta_{\ell^*+2}}{\delta_k}\right)^d \prod_{m=\ell^*+3}^{k} p_m.$$

We leave the reader to verify, along the same lines as in the previous case, that this last quantity is bounded from below by $r^{d(1-\varepsilon)}/\Phi(1/r)$, completing the proof. □

11. Appendix: A generalized Young-type inequality

Proposition 11.1. *Let μ be a positive Borel measure supported on a set $E \subseteq \mathbb{R}^d$ that is not necessarily translation-invariant. Given a measurable function $K(\cdot,\cdot)$, consider the integral operator:*

$$Tf(x) := \int K(x,y) f(y)\, d\mu(y). \tag{11.1}$$

Then for any choice of exponents $1 \leq s, q, r \leq \infty$ satisfying

$$1 + \frac{1}{r} = \frac{1}{s} + \frac{1}{q}, \tag{11.2}$$

$$A_s := \sup_{x\in E}\left[\int |K(x,y)|^s\, d\mu(y)\right]^{\frac{1}{s}} < \infty, \quad B_s := \sup_{y\in E}\left[\int |K(x,y)|^s\, d\mu(x)\right]^{\frac{1}{s}} < \infty, \tag{11.3}$$

the following inequality holds:

$$||Tf||_{L^r(\mu)} \leq A_s^{1-\frac{s}{r}} B_s^{\frac{s}{r}} ||f||_{L^q(\mu)}.$$

Proof. We adapt the same method of proof as the classical Young's inequality, where μ is the Lebesgue measure, with the modifications needed to deal with the lack of translation-invariance. We begin with a pointwise bound for the operator T.

$$\begin{aligned}
|Tf(x)| &\leq \int |K(x,y)||f(y)|\, d\mu(y) \\
&\leq \int |K(x,y)|^{1+\frac{s}{r}-\frac{s}{r}}|f(y)|^{1+\frac{q}{r}-\frac{q}{r}}\, d\mu(y) \\
&= \int \left(|K(x,y)|^{\frac{s}{r}}|f(y)|^{\frac{q}{r}}\right)|K(x,y)|^{\frac{r-s}{r}}|f(y)|^{\frac{r-q}{r}}\, d\mu(y) \\
&\leq \mathfrak{T}_1(x)\times\mathfrak{T}_2(x)\times\mathfrak{T}_3, \text{ where} \\
\mathfrak{T}_1(x) &= ||\left(|K(x,\cdot)|^s|f(\cdot)|^q\right)^{\frac{1}{r}}||_{L^r(\mu)} = \left(\int |K(x,y)|^s|f(y)|^q\, d\mu(y)\right)^{\frac{1}{r}}, \\
\mathfrak{T}_2(x) &= \||K(x,\cdot)|^{\frac{r-s}{r}}\|_{L^{\frac{sr}{r-s}}(\mu)} = \left(\int |K(x,y)|^s d\mu(y)\right)^{\frac{r-s}{sr}} \leq A_s^{\frac{r-s}{r}} \\
\mathfrak{T}_3 &= \||f|^{\frac{r-q}{r}}\|_{L^{\frac{qr}{r-q}}(\mu)} = ||f||^{\frac{r-q}{r}}_{L^q(\mu)}.
\end{aligned}$$

The last line in the estimation of $Tf(x)$ above uses the generalized Hölder's Inequality with the triple of exponents $(r, sr/(r-s), qr/(r-q))$, whose reciprocals add up to one, by (11.2). We proceed to compute the $L^r(\mu)$-norm of our convolution:

$$\begin{aligned}
\|Tf\|^r_{L^r(\mu)} &\leq \int [\mathfrak{T}_1(x)\times\mathfrak{T}_2(x)\times\mathfrak{T}_3]^r\, d\mu(x) \\
&\leq \int \left(\int |K(x,y)|^s|f(y)|^q\, d\mu(y)\right)A_s^{r-s}\times ||f||^{r-q}_{L^q(\mu)}\, d\mu(x) \\
&\leq A_s^{r-s}||f||^{r-q}_{L^q(\mu)}\iint |K(x,y)|^s|f(y)|^q\, d\mu(y)\, d\mu(x) \\
&\leq A_s^{r-s}||f||^{r-q}_{L^q(\mu)}\times (B_s^s||f||^q_{L^q(\mu)}) \\
&\leq A_s^{r-s}B_s^s||f||^r_{L^q(\mu)},
\end{aligned}$$

which completes our proof. □

12. Appendix: Probabilistic tools

12.1. **Proof of Lemma 2.3.**

Proof. The inequality (2.19) is a consequence of (2.18). We start by proving this implication. Define the auxiliary quantity

$$\eta_k := \frac{P_k - \overline{P}_k}{\overline{P}_k}, \quad \text{so that} \quad P_k = \overline{P}_k(1+\eta_k).$$

An iteration involving (2.16) then gives

$$P_k = \overline{P}_k(1+\eta_k) = N_k^{d(1-\varepsilon_k)}(1+\eta_k)P_{k-1} = \cdots$$

$$= \prod_{j=1}^{k} \big[N_j^{d(1-\varepsilon_j)} (1+\eta_j) \big] = R_k \prod_{j=1}^{k} (1+\eta_j).$$

Thus, in order to prove (2.19) it suffices to establish the $\mathbb{P}^*$-almost sure existence of a constant $C_2 = C_2(\omega) > 0$ such that

$$C_2^{-1} \le \prod_{j=1}^{k} (1+\eta_j) \le C_2 \quad \text{for all } k \ge 1. \tag{12.1}$$

This follows from two observations: the first is that for $\mathbb{P}^*$-almost every random set E,

$$\eta_k \ne -1, \text{ since } P_k \ne 0 \text{ for all } k \ge 1, \tag{12.2}$$

so the product in (12.1) is strictly positive. The second point to note is that the sequence η_k is absolutely summable. To see this, we estimate the sum as follows,

$$\begin{aligned}
\sum_{k=1}^{\infty} |\eta_k| = \frac{|P_k - \overline{P}_k|}{\overline{P}_k} &\le C_1 \sum_k \sqrt{\log(k+1)}\ \max\big(\overline{P}_k, \log(k+1)\big)^{\frac{1}{2}} \times (\overline{P}_k)^{-1} \\
&\le C_1 \left[\sum_{k:\overline{P}_k \le \log(k+1)} \frac{\log(k+1)}{\overline{P}_k} + \sum_{k:\overline{P}_k > \log(k+1)} \frac{\sqrt{\log(k+1)}}{\overline{P}_k^{1/2}} \right] \\
&\le C_1 \sum_k \frac{\log(k+1)}{\overline{P}_k^{1/2}} \\
&\le C_1 \sum_k \log(k+1)\ \big[N_k^{d(1-\varepsilon_k)} P_{k-1} \big]^{-\frac{1}{2}} \\
&\le C_1 \sum_k \log(k+1)\ N_k^{-d(1-\varepsilon_k)/2}.
\end{aligned}$$

The first inequality above follows from (2.18). In the second step, we have rewritten the sum in two parts, depending on the relative sizes of $\overline{P}_k$ and $\log(k+1)$. The penultimate inequality makes use of the defining identity of $\overline{P}_k$ in (2.16). The last inequality is a consequence of the fact that $P_{k-1} \ge 1$ on the support of $\mathbb{P}^*$. Our summability hypothesis (2.17) therefore implies that $\mathbb{P}^*$-almost surely there exists a large integer $k_0 = k_0(\omega) > 0$ depending on C_1 such that

$$\frac{1}{2} \le \prod_{j=k'}^{k} (1+\eta_j) \le 2 \quad \text{for all } k \ge k' \ge k_0.$$

Set $C_2 \ge 1$ to be any constant such that $C_2 = C_2(\omega) \ge \max\left(2C_3, 1/(2C_4)\right)$, where

$$C_3 = \sup_{1 \le k \le k_0} \prod_{j=1}^{k} (1+\eta_j) \quad \text{and} \quad C_4 = \sup_{1 \le k \le k_0} \prod_{j=1}^{k} (1+\eta_j)^{-1}.$$

Both C_3 and C_4 are strictly positive, by (12.2). Then (12.1) holds with this C_2, proving (2.19).

It remains to prove (2.18). For a fixed large ω-independent constant B soon to be specified ($B = 100$ will suffice), we define the event

$$\mathtt{S}_k := \left\{\omega \in \Omega : |P_k - \overline{P}_k| \le t_0\right\},$$

where

$$t_0 := B\sqrt{\log(k+1)}\max\Big(\overline{P}_k, \log(k+1)\Big)^{\frac{1}{2}} = \begin{cases} B\log(k+1) & \text{for } \overline{P}_k \le \log(k+1), \\ B\log(k+1)\overline{P}_k^{\frac{1}{2}} & \text{for } P_k > \log(k+1). \end{cases} \tag{12.3}$$

We aim to show that

$$\sum_{k=1}^{\infty} \mathbb{P}^*(\mathtt{S}_k^c) < \infty. \tag{12.4}$$

Once (12.4) is established, the Borel-Cantelli Lemma [2, p 53, Theorem 4.3] implies that for $\mathbb{P}^*$-almost every ω, there exists an integer $k_0(\omega) \ge 1$ such that the event $\mathtt{S}_k$ occurs for all $k \ge k_0(\omega)$. Since we have

$$\sup_{k \le k_0(\omega)} \frac{|P_k - \overline{P}_k|}{\sqrt{\overline{P}_k \log(k+1)}} \le \frac{1}{\sqrt{\log 2}} \sup_{k \le k_0(\omega)} |P_k - \overline{P}_k| \le \frac{2M_{k_0}^d}{\sqrt{\log 2}}$$

for every such ω, the desired inequality (2.18) holds $\mathbb{P}^*$-almost surely by setting $C_1 = \max(B, 2M_{k_0}^d/\sqrt{\log 2})$.

We now turn our attention to proving (12.4). Let $\mathcal{F}_k$ denote the σ-algebra generated by $\mathbf{Y}_1, \cdots, \mathbf{Y}_k$. Since

$$\mathbb{P}^*(\mathtt{S}_k^c) = \frac{\mathbb{P}(\mathtt{T}_k)}{\mathbb{P}(E \ne \emptyset)} = \frac{\mathbb{E}(\mathbb{P}(\mathtt{T}_k|\mathcal{F}_{k-1}))}{\mathbb{P}(E \ne \emptyset)} \quad \text{with } \mathtt{T}_k = \mathtt{S}_k^c \cap \{E \ne \emptyset\},$$

it suffices to show that $\mathbb{P}(\mathtt{T}_k|\mathcal{F}_{k-1})$ is bounded above by a deterministic constant that is summable in k. We estimate $\mathbb{P}(\mathtt{T}_k|\mathcal{F}_{k-1})$ using Bernstein's inequality, quoted in Theorem 12.3 below. Conditioning on $\mathcal{F}_{k-1}$, we observe that

$$P_k - \overline{P}_k = \sum_{\mathbf{i}_{k-1}} X_{k-1}(\mathbf{i}_{k-1}) \sum_{i_k=1}^{N_k^d} \left(Y_k(\mathbf{i}_k) - N_k^{-d\varepsilon_k}\right)$$

is the sum of $\overline{P}_k = P_{k-1}N_k^d$ independent, centred random variables, each of which is bounded above by 1 in absolute value and has variance $\le p_k = N_k^{-d\varepsilon_k}$. Thus in the notation of Theorem 12.3, $m = \overline{P}_k$, $M = 1$ and $\sum \sigma_r^2 \le \sigma^2 = \overline{P}_k$. We apply (12.16) with these values and with $t = t_0$ as in (12.3). This yields

$$\mathbb{P}(\mathtt{T}_k|\mathcal{F}_{k-1}) = \mathbb{P}(|P_k - \overline{P}_k| > t_0|\mathcal{F}_{k-1}) \le \exp\left(-\frac{t_0^2}{\overline{P}_k + \frac{t_0}{3}}\right) \le \begin{cases} \exp\left(-\frac{t_0^2}{4t_0/3}\right) & \text{for } \overline{P}_k \le \log(k+1), \\ \exp\left(-\frac{t_0^2}{4B\overline{P}_k/3}\right) & \text{for } P_k > \log(k+1). \end{cases}$$

$$\leq (k+1)^{-\frac{3B}{4}}.$$

The final step is obtained by substituting t_0 from (12.3) into the two cases. The right hand side is summable for any choice of $B > 4$, completing the proof. □

12.2. **Estimating the number of descendants of a basic cube.** For indices $1 \leq r < \ell$, and a fixed $\mathbf{i}_r \in \mathbb{I}(r,d)$, let us define

$$q_\ell[\mathbf{i}_r] := \sum{}' X_\ell(\mathbf{i}_\ell), \tag{12.5}$$

where $\sum'$ ranges over all multi-indices $\mathbf{i}_\ell \in \mathbb{I}(\ell,d)$ whose projection onto the first r coordinates yields $\mathbf{i}_r$. Thus $q_\ell[\mathbf{i}_r]$ represents the number of basic cubes of the k-th generation descended from $Q(\mathbf{i}_r)$.

Lemma 12.1. *Suppose that the construction parameters N_k and ε_k obey the following summability condition:*

$$\sum_{k>k'} k|\log \delta_{k'}| N_k^{-d(1-\varepsilon_k)/2} < \infty. \tag{12.6}$$

Then for $\mathbb{P}^*$-almost every $\omega \in \Omega$, there exists a constant $C = C_\omega > 0$ such that for every choice of indices $r < \ell$,

$$\sup\big\{q_\ell[\mathbf{i}_r] : \mathbf{i}_r \in \mathbb{I}(r,d)\big\} \leq C\left(\frac{\delta_r}{\delta_\ell}\right)^d \prod_{m=r+1}^{\ell} p_m.$$

Remark: We observe that the condition (12.5) is stronger that the other two conditions (2.14) and (2.17) needed in earlier parts of the argument. Further, the choices of N_k and ε_k as given in (2.20) and (2.21) satisfy (12.6) and hence also (2.14) and (2.17).

Proof. The proof is very similar to that of Lemma 2.3 above, so we only sketch the details. For fixed r, we define a partially averaged version of q_ℓ, which we call $\overline{q}_\ell$:

$$\overline{q}_\ell[\mathbf{i}_r] = N_\ell^d p_\ell q_{\ell-1}[\mathbf{i}_r] = N_\ell^{d(1-\varepsilon_\ell)} q_{\ell-1}[\mathbf{i}_r], \tag{12.7}$$

with p_ℓ and ε_ℓ as in (2.13).

For a fixed large ω-independent constant B to be specified, we set

$$\begin{aligned}\tau_0 &:= (\ell+B)^{1/2}|\log\delta_r|^{1/2}\max\big(\overline{q}_\ell[\mathbf{i}_r], (\ell+B)|\log\delta_r|\big)^{1/2}\\ &= \left\{\begin{array}{ll}(\ell+B)^{\frac{1}{2}}|\log\delta_r|^{1/2}(\overline{q}_\ell[\mathbf{i}_r])^{1/2} & \text{if } \overline{q}_\ell \geq (\ell+B)|\log\delta_r|\\ (\ell+B)|\log\delta_r| & \text{if } \overline{q}_\ell \leq (\ell+B)|\log\delta_r|,\end{array}\right\}\end{aligned} \tag{12.8}$$

and define the event

$$\mathtt{T}_r := \bigcap_{\ell=r+1}^{\infty} \mathtt{T}_{r\ell}, \quad \text{where} \quad \mathtt{T}_{r\ell} := \big\{\omega \in \Omega : |q_\ell[\mathbf{i}_r] - \overline{q}_\ell[\mathbf{i}_r]| \leq \tau_0 \text{ for all } \mathbf{i}_r \in \mathbb{I}(r,d)\big\}.$$

We aim to show that

$$\sum_{r=1}^{\infty} \mathbb{P}^*(\mathtt{T}_r^c) = \sum_{r=1}^{\infty} \mathbb{P}^*\Big(\bigcup_{\ell=r+1}^{\infty} \mathtt{T}_{r\ell}^c\Big) \leq \sum_{r=1}^{\infty}\sum_{\ell=r+1}^{\infty} \mathbb{P}^*(\mathtt{T}_{r\ell}^c) < \infty. \tag{12.9}$$

The same Borel-Cantelli argument as in Lemma 2.3 would then imply that for almost every $\omega \in \Omega$, there is a constant $C > 0$ such that

$$|q_\ell[\mathbf{i}_r] - \overline{q}_\ell[\mathbf{i}_r]| \leq C\tau_0 \text{ for all } r < \ell \text{ and all } \mathbf{i}_r. \tag{12.10}$$

We will return to the proof of (12.9) shortly, but will leave the verification of (12.10) from (12.9) to the reader. Assuming (12.10) for the moment, the remainder of the proof is completed as follows. We fix indices $r < \ell$, and a multi-index $\mathbf{i}_r$, and for simplicity write $q_\ell = q_\ell[\mathbf{i}_r]$, $\overline{q}_\ell = \overline{q}_\ell[\mathbf{i}_r]$. Define an auxiliary quantity ζ_ℓ according to the following relation

$$q_\ell = \overline{q}_\ell(1 + \zeta_\ell). \tag{12.11}$$

The definitions (12.5) and (12.7) of q_ℓ and $\overline{q}_\ell$ imply if q_ℓ is nonzero, then so is every q_m and $\overline{q}_m$ for $r + 1 \leq m \leq \ell$. As a result, ζ_ℓ is well-defined for nonzero q_ℓ. Further, (12.10) implies that in this case,

$$\begin{aligned} |\zeta_\ell| \leq \frac{C\tau_0}{\overline{q}_\ell} &\leq C(\ell + B)|\log \delta_r| \times (\overline{q}_\ell)^{-1/2} \\ &\leq C(\ell + B)|\log \delta_r| N_\ell^{-d(1-\varepsilon_\ell)/2} \end{aligned} \tag{12.12}$$

The second inequality in the sequence above follows from (12.8). The last inequality is a consequence of (12.7), which says that if $\overline{q}_\ell$ is nonzero, it must be larger than $N_\ell^d p_\ell$.

The quantity ζ_ℓ is analogous to η_k in the proof of Lemma 2.3 and will play a similar role. Iterating the relation (12.11) and applying (12.7) at every step, we arrive at

$$\begin{aligned} q_\ell &= N_\ell^d p_\ell q_{\ell-1}(1 + \zeta_\ell) = \cdots \\ &= \Big[\prod_{m=r+1}^{\ell} N_m^d p_m\Big]\Big[\prod_{m=r+1}^{\ell}(1 + \zeta_m)\Big] q_r = \Big(\frac{\delta_r}{\delta_\ell}\Big)^d \Big[\prod_{m=r+1}^{\ell} p_m\Big]\Big[\prod_{m=r+1}^{\ell}(1 + \zeta_m)\Big]. \end{aligned} \tag{12.13}$$

Since $q_r = q_r[\mathbf{i}_r] = 1$, it suffices to show that the second product above is bounded above by a constant independent of r and ℓ and depending only on ω. The estimate (12.12) shows that ζ_m is small for $m \geq r$ and large r, so in order to establish the desired conclusion it suffices to show that the sum of $|\zeta_m|$ in the range $r + 1 \leq m \leq \ell$ is bounded above by a large ω-dependent constant that is uniform in r and ℓ. The summability hypothesis (12.6) ensures that this is the case.

It remains to prove (12.9). We do this again with an application of Bernstein's inequality, as we did in the proof of Lemma 2.3. Since

$$q_\ell - \overline{q}_\ell = \sum_{\mathbf{i}_{\ell-1}}{}' X_{\ell-1}(\mathbf{i}_{\ell-1}) \sum_{\bar{i}_\ell} \big(Y_\ell(\mathbf{i}_\ell) - p_\ell\big),$$

we can set

$$m = q_{\ell-1} N_\ell^d, \qquad \sigma^2 = N_\ell^d q_{\ell-1} p_\ell = \overline{q}_\ell \qquad \text{and} \qquad t = \tau_0$$

in Theorem 12.3 to deduce that for each $\mathbf{i}_r \in \mathbb{I}(r, d)$,

$$\mathbb{P}^*\big(|q_\ell - \overline{q}_\ell| > \tau_0\big) \leq \exp\left(-\frac{\tau_0^2}{\overline{q}_\ell + \frac{\tau_0}{3}}\right) \leq \delta_r^{\frac{3}{4}(\ell+B)}.$$

Since the number of possible choices of multi-indices $\mathbf{i}_r$ is δ_r^{-d}, summing the above estimate over all $\mathbf{i}_r$ yields

$$\mathbb{P}^*(\mathtt{T}_{r\ell}^c) \leq \delta_r^{\frac{3}{4}(\ell+B)-d}.$$

Choosing $B > 4d/3$ ensures that the last quantity is summable in r and ℓ for all $\ell \geq r+1$. This completes the proof of (12.9) and hence the proof of the lemma. □

A careful analysis of the proof of Lemma 12.1 yields a lower bound on q_ℓ as well.

Lemma 12.2. *Assume that the summability condition* (12.6)*. Then for* $\mathbb{P}^*$*-a.e.* $\omega \in \Omega$ *and every* $x \in E = E(\omega)$*, the following property holds:*

There exists a constant $C = C_{x,\omega} > 0$ *such that for all indices* $r < \ell$ *and all multi-indices* $\mathbf{i}_r$ *and* $\mathbf{i}_\ell$ *such that*

$$x \in Q(\mathbf{i}_\ell) \subseteq Q(\mathbf{i}_r), \tag{12.14}$$

we have

$$q_\ell[\mathbf{i}_r] \geq C^{-1}\left(\frac{\delta_r}{\delta_\ell}\right)^d \prod_{m=r+1}^{\ell} p_m. \tag{12.15}$$

Proof. We proceed exactly as in Lemma 12.1, leading up to the relation (12.13). The hypothesis (12.14) implies that q_ℓ is nonzero for all the relevant choices of $\mathbf{i}_r$ and $\mathbf{i}_\ell$; in particular, none of the factors $1+\zeta_m$ can be zero, for $r+1 \leq m \leq \ell$. On the other hand, the estimate (12.12) and the summability hypothesis (12.6) imply that the tail product of $\prod_m(1+\zeta_m)$ converges to a nonzero quantity, i.e., there are large absolute constants $R, C > 0$ such that

$$\inf\Big\{\prod_{m=r+1}^{\ell}(1+\zeta_m) : \ell > r \geq R\Big\} \geq C^{-1}.$$

This leaves at most R factors unaccounted for, but since each factor is nonzero, we can reach (12.15) simply by enlarging C by a constant factor depending only on R, x and ω. □

12.3. **Large deviation inequalities.**

Theorem 12.3. *(Bernstein's inequality) Let* $Z_1, \dots, Z_m$ *be independent random variables with*

$$|Z_r| \leq M, \quad \mathbb{E}Z_r = 0 \quad \text{and} \quad \mathbb{E}|Z_r|^2 = \sigma_r^2.$$

Let $\sum \sigma_r^2 \leq \sigma^2$*. Then for all* $t > 0$*,*

$$\mathbb{P}\Big(\big|\sum_1^m Z_r\big| \geq t\Big) \leq \exp\left(-\frac{t^2}{\sigma^2 + \frac{1}{3}Mt}\right). \tag{12.16}$$

Theorem 12.4. *(Azuma's inequality) Suppose that* $\{U_r : r = 0, 1, 2, \dots\}$ *is a martingale and* $\{c_r : r \geq 0\}$ *is a sequence of positive numbers such that* $|U_{r+1} - U_r| \leq c_r$ *almost surely. Then for all integers* $m \geq 1$ *and all* $t \in \mathbb{R}$*,*

$$\mathbb{P}\left(|U_m - U_0| \geq t\right) \leq 2\exp\Big[-t^2/\big(2\sum_{r=1}^m c_r^2\big)\Big] \tag{12.17}$$

13. Acknowledgements

This work was partially done while SE was on sabbatical at McGill University. He warmly thanks the Department of Mathematics for their hospitality and working environment. A Heilbronn Institute 10th Anniversary Grant funded SE during the initial stages of this project in 2016. MP was supported by a 2018 Wall Scholarship from the Peter Wall Institute of Advanced Study and an NSERC Discovery grant.

References

[1] N. Anantharaman, *Entropy and the localization of eigenfunctions*, Ann. of Math. (2) **168** no. 2, 435-475 (2008).

[2] P. Billingsley, *Probability and measure*, Second edition, 1991, Wiley series in Probability and Mathematical Statistics.

[3] M. Blair and C. Sogge, *Logarithmic improvements in L^p-norms for eigenfunctions at the critical exponent in the presence of nonpositive curvature*, preprint, arxiv.org/abs/1706.06704 (2018).

[4] J. Bourgain, *Geodesic restriction and L^p-estimates for eigenfunctions on Riemannian surfaces*, Linear and complex analysis, 2735, Amer. Math. Soc. Transl. Ser. 2, 226, Adv. Math. Sci., 63 (2009).

[5] J. Bourgain and S. Dyatlov, *Fourier dimension and spectral gaps for hyperbolic surfaces*, Geom. Funct. Anal. **27**, 744-771 (2017).

[6] J. Bourgain and S. Dyatlov, *Spectral gaps without the pressure condition*, Ann. Math. **187** no. 3, 824-867 (2018).

[7] J. Bourgain and Z. Rudnick, *Restriction of toral eigenfunctions to hypersurfaces and nodal sets*, Geom. Funct. Anal. **22**, 878-937 (2012).

[8] N. Burq, P. Gérard, N. Tzetkov, *Restrictions of the Laplace-Beltrami eigenfunctions to submanifolds*, Duke Math. J. **138** no. 3, 445-487 (2007).

[9] Y. Colin de Verdière, *Ergodicite et fonctions propres du laplacien*, Comm. Math. Phys. **102**, 497-502 (1985).

[10] X. Chen, C. Sogge, *A few endpoint geodesic restriction estimates for eigenfunctions*, Comm. Math. Phys. **329**, 435-459 (2014).

[11] K. Falconer, *Fractal geometry: Mathematical foundations and applications*, First edition, Wiley Brothers (1990).

[12] P. Gérard and É. Leichtnam, *Ergodic properties of eigenfunctions for the Dirichlet problem*, Duke Math. J. **71**, 559-607 (1993).

[13] A. Hassell and M. Tacy, *Semiclassical Lp estimates of quasimodes on curved hypersurfaces*, J. Geom. Anal. **22**, 74-89 (2012).

[14] B. Helffer, A. Martinez, and D. Robert, *Ergodicité et limite semi-classique*, Comm. Math. Phys. **109**, 313-326 (1987).

[15] H. Hezari and G. Riviére, *Equidistribution of toral eigenfunctions along hypersurfaces*, to appear in Rev. Iber. Math. (2018).

[16] L. Hörmander, *Spectral function of an elliptic operator*, Acta Math., **121**, 193-218 (1968).

[17] L. Hörmander, *Oscillatory integrals and multipliers on FL^p*, Ark. Math. **11**, 1-11 (1971).

[18] R. Hu, *L^p norm estimates of eigenfunctions restricted to submanifolds*, Forum Math., **21**, 1021-1052 (2009).

[19] J.-P. Kahane and J. Peyriere, *Sur certaines martingales de Benoît Mandelbrot*, Adv. Math., **22** (2), 131-145 (1976).

[20] H. Koch, D. Tataru, and M. Zworski, *Semiclassical L^p estimates*, Ann. Henri Poincare, **5**, 885-916 (2007).

[21] I. Laba and M. Pramanik, *Arithmetic progressions in sets of fractal dimensions*, Geom. Funct. Anal., **19**, 429-456 (2009).

[22] I. Laba and M. Pramanik, *Maximal operators and differentiation theorems for sparse sets*, Duke Math. J., **158**, no. 3, 347-411 (2011).
[23] E. Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity*, Ann. of Math. (2) **163**, 165-219 (2006).
[24] P. Mattila, *Geometry of sets and measures in Euclidean space*, Cambridge studies in advanced mathematics, Cambridge Univ. Press, Cambridge (2012).
[25] F. Naud, *Expanding maps on Cantor sets and analytic continuation of zeta functions*, Ann. de l'ENS **4** 38, 116-153 (2005).
[26] A. Reznikov, *Norms of geodesic restrictions for eigenfunctions on hyperbolic surfaces and representation theory*, unpublished update to a work from 2005, arxiv.org/abs/math/0403437v3 (2010).
[27] Z. Rudnick and P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** no. 1, 195-213 (1994).
[28] P. Sarnak, "Arithmetic quantum chaos" in *The Schur lectures (Tel Aviv, 1992)*, Israel Math. Conf. Proc. **8**, Bar-Ilan Univ., Ramat Gan, Israel, 183-236 (1995).
[29] P. Shmerkin and V. Suomala, *A class of random Cantor measures, with applications*, Recent developments in fractals and related fields, 233260, Trends Math., Birkhuser/Springer, Cham (2017).
[30] P. Shmerkin and V. Suomala, *Spatially independent martingales, intersections, and applications*, Mem. Amer. Math. Soc. **251**, no. 1195 (2018).
[31] A. I. Shnirelman, *Ergodic properties of eigenfunctions* (in Russian), Uspekhi Mat. Nauk **29**, no. 6, 181-182 (1974).
[32] C. Sogge, *Oscillatory integrals and spherical harmonics*, Duke Math. J. **53**, 43-65 (1986).
[33] C. Sogge, *Concerning the L^p norms of spectral clusters of second-order elliptic operators on compact manifolds*, J. Func. Analysis, **77**, 123-138 (1988).
[34] C. Sogge, *Fourier Integrals in Classical Analysis*, Cambridge Tracts in Math., Cambridge Univ. Press, Cambridge (1993).
[35] M. Tacy, *Semiclassical L^p estimates of quasimodes on submanifolds"*, Comm. PDE [35], no. 8, 1538-1562 (2010).
[36] P. Tomas, *A restriction theorem for the Fourier transform*, Bull. Amer. Math. Soc. **81**, 477-478 (1975).
[37] P. Yung, *Spectral projection theorems on compact manifolds*, Lecture notes, available under "Notes/Slides" at https://www.math.cuhk.edu.hk/~plyung/Sogge.pdf.
[38] S. Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*, Duke Math. J. **55**, 919-941 (1987).
[39] S. Zelditch and M. Zworski, *Ergodicity of eigenfunctions for ergodic billiards*, Comm. Math. Phys. **175**, 673-682 (1996).

(Permanent) School of Mathematics, Cardiff University, Senghennyd Road, Cardiff, Wales, United Kingdom
E-mail address: eswarathasans@cardiff.ac.uk

(Current) Department of Mathematics and Statistics, McGill University, Burnside Hall, Montréal, QC, Canada
E-mail address: suresh.eswarathasan@mcgill.ca

University of British Columbia, Department of Mathematics, 1984 Mathematics Road, Vancouver, BC V6T 1Z2
E-mail address: malabika@math.ubc.ca